09/673274

```
                  GenCore version 5.1.6
         Copyright (c) 1993 - 2003  Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on:        October 16, 2003, 17:06:33 ; Search time 8.77076 Seconds
                                            (without alignments)
                                            1415.50 Million cell updates/sec

Title:         US-09-673-274B-2_COPY_1_132
Perfect score: 267
Sequence:      1 gataagcttatcggttcctg..........tgaactgctggtgcgagact 132

Scoring table: BLOSUM62
               Xgapop 10.0 , Xgapext  0.5
               Ygapop 10.0 , Ygapext  0.5
               Fgapop  6.0 , Fgapext  7.0
               Delop   6.0 , Delext   7.0

Searched:      127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters:      255726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09673274/runat_16102003_170627_2531/app_query.fasta_1.654
-DB=SwissProt_41 -QFMT=fastan -SUFFIX=rsp -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09673274_@CGN_1_1_36_@runat_16102003_170627_2531 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :        SwissProt_41:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
                  %
Result          Query
  No.   Score  Match Length DB ID                         Description
-------------------------------------------------------------------------
    1     267  100.0     44  1 DEFN_HELVI                 P81544 heliothis v
    2      86   32.2     66  1 SCX3_BUTOM                 P13488 buthus occi
    3    84.5   31.6     85  1 SCN2_MESMA                 Q9bkj1 mesobuthus
    4      84   31.5     85  1 SCAS_MESMA                 Q9uac9 mesobuthus
    5    83.5   31.3     85  1 SCN3_MESMA                 Q9bkj0 mesobuthus
    6    79.5   29.8     66  1 SCXC_CENLL                 P45667 centruroide
    7      79   29.6     85  1 SCXE_BUTOC                 Q17254 buthus occi
    8    78.5   29.4     65  1 SCX4_BUTOM                 P59354 buthus occi
    9    78.5   29.4     84  1 SCX3_ANDAU                 P01480 androctonus
   10    77.5   29.0     64  1 SCL6_LEIQH                 P59356 leiurus qui
   11    77.5   29.0     65  1 SCX4_LEIQU                 P01489 leiurus qui
   12      77   28.8     85  1 SCAT_MESMA                 Q9uac8 mesobuthus
   13    76.5   28.7     61  1 SIX2_LEIQU                 P19855 leiurus qui
   14      76   28.5     66  1 SCXE_BUTEU                 P09982 buthus eupe
   15      76   28.5     67  1 SCL3_LEIQH                 P56678 leiurus qui
   16    75.5   28.3     84  1 SCXA_MESMA                 O61705 mesobuthus
   17    74.5   27.9     70  1 DMYC_DROME                 P41964 drosophila
   18    74.5   27.9     87  1 SCX1_CENSC                 P01492 centruroide
   19      74   27.7     59  1 SCX5_CENSC                 P58779 centruroide
   20      74   27.7     84  1 SCX4_ANDAU                 P45658 androctonus
   21      73   27.3     61  1 SIX5_ANDAU                 P81504 androctonus
   22      73   27.3     64  1 SCX5_LEIQU                 P01481 leiurus qui
   23      72   27.0     85  1 SC15_MESMA                 Q9gng8 mesobuthus
   24      72   27.0     85  1 SCA1_MESMA                 Q9gyx2 mesobuthus
   25    71.5   26.8     61  1 SIX2_BUTAR                 P80962 buthacus ar
   26    71.5   26.8     61  1 SIX4_BUTOC                 P55903 buthus occi
   27    71.5   26.8     61  1 SIX5_BUTOC                 P55904 buthus occi
   28    71.5   26.8     87  1 SCX2_CENSC                 P01493 centruroide
   29    71.5   26.8     87  1 SCX3_CENSC                 P01494 centruroide
   30      71   26.6     43  1 DEF1_PALPR                 P80407 palomena pr
   31      71   26.6     66  1 SCL7_LEIQH                 P59357 leiurus qui
   32    70.5   26.4     64  1 SCM2_MESMA                 P58488 mesobuthus
   33    70.5   26.4     84  1 SC16_MESMA                 Q9gqv6 mesobuthus
   34    70.5   26.4     87  1 SCX6_CENNO                 P45664 centruroide
   35    69.5   26.0     84  1 SCX1_MESMA                 P45697 mesobuthus
   36    69.5   26.0     84  1 SCX2_CENNO                 P01495 centruroide
   37    69.5   26.0     84  1 SCX9_CENSC                 Q95wc9 centruroide
   38    69.5   26.0     85  1 SC10_MESMA                 Q9njc5 mesobuthus
   39    68.5   25.7     62  1 SIX4_BUTSI                 P82814 buthus sind
   40    68.5   25.7     65  1 SCX1_BUTOC                 P01488 buthus occi
   41    68.5   25.7     66  1 SCX6_ANDAU                 P56743 androctonus
   42    68.5   25.7     66  1 SCX9_BUTEU                 P09981 buthus eupe
   43    68.5   25.7     74  1 DEF2_CAPAN                 O65740 capsicum an
   44    68.5   25.7     74  1 SC17_MESMA                 Q9njc4 mesobuthus
   45    68.5   25.7     83  1 SCX1_ANDAU                 P01479 androctonus


                              ALIGNMENTS

RESULT 1
DEFN_HELVI
ID   DEFN_HELVI     STANDARD;      PRT;    44 AA.
AC   P81544;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Defensin heliomicin.
OS   Heliothis virescens (Noctuid moth) (Owlet moth).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea;
OC   Noctuidae; Heliothinae; Heliothis.
OX   NCBI_TaxID=7102;
RN   [1]
RP   SEQUENCE, AND CHARACTERIZATION.
RC   TISSUE=Hemolymph;
RX   MEDLINE=99194775; PubMed=10092609;
RA   Lamberty M., Ades S., Uttenweiler-Joseph S., Brookhart G., Bushey D.,
RA   Hoffmann J.A., Bulet P.;
RT   "Insect immunity. Isolation from the lepidopteran Heliothis virescens
RT   of a novel insect defensin with potent antifungal activity.";
RL   J. Biol. Chem. 274:9320-9326(1999).
CC   -!- FUNCTION: THIS PEPTIDE HAS POTENT ANTI-FUNGAL ACTIVITY. HAS NO
CC       ACTIVITY AGAINST GRAM-NEGATIVE AND GRAM-POSITIVE BACTERIA.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- MASS SPECTROMETRY: MW=4784.8; METHOD=MALDI.
DR   PDB; 1I2U; 12-FEB-02.
DR   PDB; 1I2V; 12-FEB-02.
KW   Insect immunity; Fungicide; 3D-structure.
FT   DISULFID      7     32
FT   DISULFID     18     40
FT   DISULFID     22     42
SQ   SEQUENCE   44 AA;  4790 MW;  301DA1BD3376FCD5 CRC64;

Alignment Scores:
Pred. No.:              3.22e-27      Length:        44
Score:                  267.00        Matches:       44
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            100.00%       Indels:        0
DB:                     1             Gaps:          0
```

09/673274

```
US-09-673-274B-2_COPY_1_132 (1-132) x DEFN_HELVI (1-44)

Qy          1 GATAAGCTTATCGGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
                |||:::|||..|||:|||||||::|||:::|||.:|||.:|||.||||||||||||||||
Db          1 AspLysLeuIleGlySerCysValTrpGlyAlaValAsnTyrThrSerAspCysAsnGly 20

Qy         61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
                ::|||.:||||| ||||:::||||:.|||.:||||:.|||:.|||:::|||||||||||||
Db         21 GluCysLysArgArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnValAsnCys 40

Qy        121 TGGTGCGAGACT 132
                ::|||||||:||||
Db         41 TrpCysGluThr 44

RESULT 2
SCX3_BUTOM
ID   SCX3_BUTOM      STANDARD;      PRT;    66 AA.
AC   P13488;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Alpha-like neurotoxin III (Bom III) (Bom3).
OS   Buthus occitanus mardochei (Moroccan scorpion).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Buthus.
OX   NCBI_TaxID=6869;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Venom;
RX   MEDLINE 87161829; PubMed=3104036;
RA   Vargas O., Martin M.-F., Rochat H.;
RT   "Characterization of six toxins from the venom of the Moroccan
RT   scorpion Buthus occitanus mardochei.";
RL   Eur. J. Biochem. 162:589-599(1987).
CC   -!- FUNCTION: Binds to sodium channels and inhibits the inactivation
CC       of the activated channels, thereby blocking neuronal transmission.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC       ALPHA-TOXIN SUBFAMILY. ALPHA-LIKE TOXIN SUBCLASS.
DR   PIR; A29386; A29386.
DR   HSSP; P56678; 1BMR.
DR   InterPro; IPR003614; Knot1.
DR   InterPro; IPR002061; Scorpion_toxinL.
DR   Pfam; PF00537; toxin_3; 1.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor.
FT   DISULFID     12     65       BY SIMILARITY.
FT   DISULFID     16     37       BY SIMILARITY.
FT   DISULFID     23     47       BY SIMILARITY.
FT   DISULFID     27     49       BY SIMILARITY.
SQ   SEQUENCE   66 AA;  6872 MW;  7A96DA393F89952F CRC64;

Alignment Scores:
Pred. No.:                0.00211        Length:         66
Score:                    86.00          Matches:        15
Percent Similarity:       55.00%         Conservative:   7
Best Local Similarity:    37.50%         Mismatches:     16
Query Match:              32.21%         Indels:         2
DB:                       1              Gaps:           1

US-09-673-274B-2_COPY_1_132 (1-132) x SCX3_BUTOM (1-66)

Qy         16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
                :::|.|||||:::                    :::|||   |||:::         ||||||   :::
Db         11 AsnCysValTyrHisCysPheProGlySerSerGlyCysAspThrLeuCysLysGluLys 30

Qy         76 GGTTACAAGGGTGGTCACTGCGGTTCCTTC------GCTAACGTGAACTGCTGGTGCGAG 129
                |||           |||||:|||||::              :::   |||   |:::|||.||:::
Db         31 GlyAlaThrSerGlyHisCysGlyPheLeuProGlySerGlyValAlaCysTrpCysAsp 50

RESULT 3
```

```
SCN2_MESMA
ID   SCN2_MESMA     STANDARD;      PRT;    85 AA.
AC   Q9BKJ1;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Anti-neuroexcitation peptide II precursor (ANEPII).
OS   Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Mesobuthus.
OX   NCBI_TaxID=34649;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Venom gland;
RA   Zhang J.-H., Hua Z.C., Zhu D.X.;
RT   "Cloning of anti-neuroexcitation peptide II (ANEP) cDNA from Scorpion
RT   Buthus martensii Karsch.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Binds to sodium channels and inhibits them.
CC   -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC   -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC       ALPHA-TOXIN SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF242736; AAK28341.1; -.
DR   HSSP; P01491; 1B3C.
DR   InterPro; IPR003614; Knot1.
DR   InterPro; IPR002061; Scorpion_toxinL.
DR   Pfam; PF00537; toxin_3; 1.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor;
KW   Signal.
FT   SIGNAL         1     21       POTENTIAL.
FT   CHAIN         22     85       ANTI-NEUROEXCITATION PEPTIDE II.
FT   DISULFID      31     81       BY SIMILARITY.
FT   DISULFID      35     56       BY SIMILARITY.
FT   DISULFID      42     63       BY SIMILARITY.
FT   DISULFID      46     65       BY SIMILARITY.
SQ   SEQUENCE   85 AA;  9224 MW;  C045C7DACB3513AB CRC64;

Alignment Scores:
Pred. No.:                0.00337        Length:         85
Score:                    84.50          Matches:        15
Percent Similarity:       55.26%         Conservative:   6
Best Local Similarity:    39.47%         Mismatches:     12
Query Match:              31.65%         Indels:         5
DB:                       1              Gaps:           2

US-09-673-274B-2_COPY_1_132 (1-132) x SCN2_MESMA (1-85)

Qy         16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
                ||||||:::||||||                                  ||||||   ||||||:::
Db         34 SerCysLeuTrpGly------------AsnAspGlyCysAsnLysGluCysArgAlaTyr 49

Qy         76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
                |||           |||:::|||   ::::::        :::   ||||||||||||
Db         50 GlyAlaSerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 66

RESULT 4
SCAS_MESMA
ID   SCAS_MESMA     STANDARD;      PRT;    85 AA.
AC   Q9UAC9;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
```

```
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Neurotoxin AS precursor.
OS   Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Mesobuthus.
OX   NCBI_TaxID=34649;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Venom gland;
RA   Lan Z.D., Dai L., Chi C.-W.;
RT   "The gene cloning and sequencing of AS and AS-1, two novel neurotoxins
RT   from the scorpion Buthus martensi Karsch.";
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Ryanodine receptor agonist in skeletal muscle.
CC   -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC   -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC       ALPHA-TOXIN SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF079060; AAD47374.1; -.
DR   HSSP; P01492; 1VNB.
DR   InterPro; IPR003614; Knot1.
DR   InterPro; IPR002061; Scorpion_toxinL.
DR   Pfam; PF00537; toxin_3; 1.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Toxin; Neurotoxin; Signal.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20     85       NEUROTOXIN AS.
FT   DISULFID     31     81       BY SIMILARITY.
FT   DISULFID     35     56       BY SIMILARITY.
FT   DISULFID     42     63       BY SIMILARITY.
FT   DISULFID     46     65       BY SIMILARITY.
SQ   SEQUENCE   85 AA;  9759 MW;  41286F9DDCD4FE90 CRC64;

Alignment Scores:
Pred. No.:               0.00392       Length:         85
Score:                   84.00         Matches:        19
Percent Similarity:      55.81%        Conservative:   5
Best Local Similarity:   44.19%        Mismatches:     17
Query Match:             31.46%        Indels:         2
DB:                      1             Gaps:           2

US-09-673-274B-2_COPY_1_132 (1-132) x SCAS_MESMA (1-85)

Qy          1 GATAAGCTTATCGGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
              ||||||       |||       ||||||       :::|||       |||   ||||||
Db         26 AspLysTyrThrGlyCysLysValTrpCysValIleAsnAsnGluSer---CysAsnSer 44

Qy         61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
              |||||||||    ||||||                |||:::|||        ||.        :::    |||
Db         45 GluCysLysIleArgGlyGlyTyrTyrGlyTyrCys---TyrPheTrpLysLeuAlaCys 63

Qy        121 TGGTGCGAG 129
              :::|||:::
Db         64 PheCysGln 66

RESULT 5
SCN3_MESMA
ID   SCN3_MESMA     STANDARD;      PRT;    85 AA.
AC   Q9BKJ0; Q9GZC4;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Anti-neuroexcitation peptide III precursor (ANEPIII) (Anti-epilepsy
DE   peptide).
OS   Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Mesobuthus.
OX   NCBI_TaxID=34649;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Venom gland;
RA   Zhang J.-H., Hua Z.C., Zhu D.X.;
RT   "Cloning of anti-epilepsy peptide cDNA from scorpion Buthus martensii
RT   Karsch.";
RL   Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Venom gland;
RA   Zhang J.-H., Hua Z.C., Zhu D.X.;
RT   "Cloning of anti-neuroexcitation peptide III (ANEP) cDNA from Scorpion
RT   Buthus martensii Karsch.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Binds to sodium channels and inhibits them.
CC   -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC   -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC       ALPHA-TOXIN SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF122003; AAG01571.1; -.
DR   EMBL; AF242737; AAK28342.1; -.
DR   HSSP; P01494; 2SN3.
DR   InterPro; IPR003614; Knot1.
DR   InterPro; IPR002061; Scorpion_toxinL.
DR   Pfam; PF00537; toxin_3; 1.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor;
KW   Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22     85       ANTI-NEUROEXCITATION PEPTIDE III.
FT   DISULFID     31     81       BY SIMILARITY.
FT   DISULFID     35     56       BY SIMILARITY.
FT   DISULFID     42     63       BY SIMILARITY.
FT   DISULFID     46     65       BY SIMILARITY.
FT   CONFLICT     47     47       K -> I (IN REF. 2).
SQ   SEQUENCE   85 AA;  9271 MW;  C1D6C93A2F82F8C2 CRC64;

Alignment Scores:
Pred. No.:               0.00457       Length:         85
Score:                   83.50         Matches:        16
Percent Similarity:      55.26%        Conservative:   5
Best Local Similarity:   42.11%        Mismatches:     12
Query Match:             31.27%        Indels:         5
DB:                      1             Gaps:           2

US-09-673-274B-2_COPY_1_132 (1-132) x SCN3_MESMA (1-85)

Qy         16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
              |||||||:::||||||                                    ||||||    ||.|||||
Db         34 SerCysLeuTrpGly------------AsnGluGlyCysAsnLysGluCysLysGlyPhe 49

Qy         76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
              |||           |||:::|||   ::::::       :::    ||||||  ||||
Db         50 GlyAlaTyrTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 66

RESULT 6
SCXC_CENLL
ID   SCXC_CENLL     STANDARD;      PRT;    66 AA.
```

```
AC   P45667;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Crustacean-specific toxin 1.
OS   Centruroides limpidus limpidus (Mexican scorpion).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Centruroides.
OX   NCBI_TaxID=29941;
RN   [1]
RP   SEQUENCE, DISULFIDE BONDS, AND STRUCTURE BY NMR.
RC   TISSUE=Venom;
RX   MEDLINE=95244424; PubMed=7727365;
RA   Lebreton F., Delepierre M., Ramirez A.N., Balderas C., Possani L.D.;
RT   "Primary and NMR three-dimensional structure determination of a novel
RT   crustacean toxin from the venom of the scorpion Centruroides limpidus
RT   limpidus Karsch.";
RL   Biochemistry 33:11135-11149(1994).
CC   -!- FUNCTION: Binds to sodium channels and inhibits the inactivation
CC       of the activated channels, thereby blocking neuronal transmission.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC       BETA-TOXIN SUBFAMILY.
DR   PIR; A55869; A55869.
DR   HSSP; P01492; 1VNB.
DR   InterPro; IPR003614; Knot1.
DR   InterPro; IPR001219; Neurotoxin.
DR   InterPro; IPR002061; Scorpion_toxinL.
DR   Pfam; PF00537; toxin_3; 1.
DR   PRINTS; PR00284; TOXIN.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor.
FT   DISULFID     12     65
FT   DISULFID     16     41
FT   DISULFID     25     46
FT   DISULFID     29     48
SQ   SEQUENCE   66 AA;  7338 MW;  A478E63796F54DAA CRC64;

Alignment Scores:
Pred. No.:                0.0151        Length:        66
Score:                    79.50         Matches:       17
Percent Similarity:       53.85%        Conservative:  4
Best Local Similarity:    43.59%        Mismatches:    11
Query Match:              29.78%        Indels:        7
DB:                       1             Gaps:          3

US-09-673-274B-2_COPY_1_132 (1-132) x SCXC_CENLL (1-66)

Qy          19 TGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGGGGT 78
                ||   |||         |.|         :::|||:::   || |||:|   :::
Db          16 CysPheTrpLeuGlyLysAsn------GluAsnCysAspLysGluCysLysAlaLysAsn 33

Qy          79 TACAAGGGT------GGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
                     |||      |||:::|||   :|||:|:||         |||||:||||:|
Db          34 GlnGlyGlySerTyrGlyTyrCysTyrSerPheAla---------CysTrpCysGlu 49

RESULT 7
SCXE_BUTOC
ID   SCXE_BUTOC     STANDARD;      PRT;    85 AA.
AC   Q17254;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Neurotoxin XIV precursor (Bot14).
OS   Buthus occitanus tunetanus (Common European scorpion).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Buthus.
OX   NCBI_TaxID=6871;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Venom;
RX   MEDLINE=96300228; PubMed=8706664;
RA   Bouhaouala-Zahar B., Ducancel F., Zenouaki I., Ben Khalifa R.,
RA   Borchani L., Pelhate M., Boulain J.C., el Ayeb M., Menez A.,
RA   Karoui H.;
RT   "A recombinant insect-specific alpha-toxin of Buthus occitanus
RT   tunetanus scorpion confers protection against homologous mammal
RT   toxins.";
RL   Eur. J. Biochem. 238:653-660(1996).
CC   -!- FUNCTION: Binds to sodium channels and inhibits the inactivation
CC       of the activated channels, thereby blocking neuronal transmission.
CC       This toxin is active only on insects.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC       ALPHA-TOXIN SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X92376; CAA63120.1; -.
DR   PIR; S68906; S68906.
DR   HSSP; P56678; 1BMR.
DR   InterPro; IPR003614; Knot1.
DR   InterPro; IPR002061; Scorpion_toxinL.
DR   Pfam; PF00537; toxin_3; 1.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor;
KW   Signal.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19     85       NEUROTOXIN XIV.
FT   DISULFID     30     83       BY SIMILARITY.
FT   DISULFID     34     55       BY SIMILARITY.
FT   DISULFID     41     65       BY SIMILARITY.
FT   DISULFID     45     67       BY SIMILARITY.
SQ   SEQUENCE   85 AA;  9174 MW;  4977F16C511F0567 CRC64;

Alignment Scores:
Pred. No.:                0.0178        Length:        85
Score:                    79.00         Matches:       14
Percent Similarity:       50.00%        Conservative:  6
Best Local Similarity:    35.00%        Mismatches:    18
Query Match:              29.59%        Indels:        2
DB:                       1             Gaps:          1

US-09-673-274B-2_COPY_1_132 (1-132) x SCXE_BUTOC (1-85)

Qy          16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
               :::|||   :::      :::      :::|||   |||:::      ||||||
Db          29 AsnCysAlaTyrHisCysLeuLysIleSerSerGlyCysAspThrLeuCysLysGluAsn 48

Qy          76 GGTTACAAGGGTGGTCACTGCGGT------TCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
               |||         ||||||||||||      |||               |||||||||:::
Db          49 GlyAlaThrSerGlyHisCysGlyHisLysSerGlyHisGlySerAlaCysTrpCysLys 68

RESULT 8
SCX4_BUTOM
ID   SCX4_BUTOM     STANDARD;      PRT;    65 AA.
AC   P59354;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Alpha-like neurotoxin IV (Bom IV) (Bom4).
OS   Buthus occitanus mardochei (Moroccan scorpion).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Buthus.
OX   NCBI_TaxID=6869;
RN   [1]
```

```
RP   SEQUENCE, PHARMACOLOGICAL CHARACTERIZATION, AND MASS SPECTROMETRY.
RC   TISSUE=Venom;
RX   MEDLINE=99203324; PubMed=10103091;
RA   Cestele S., Stankiewicz M., Mansuelle P., De Waard M., Dargent B.,
RA   Gilles N., Pelhate M., Rochat H., Martin-Eauclaire M.-F., Gordon D.;
RT   "Scorpion alpha-like toxins, toxic to both mammals and insects,
RT   differentially interact with receptor site 3 on voltage-gated sodium
RT   channels in mammals and insects.";
RL   Eur. J. Neurosci. 11:975-985(1999).
CC   -!- FUNCTION: Binds to sodium channels and inhibits the inactivation
CC       of the activated channels, thereby blocking neuronal transmission.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC   -!- MASS SPECTROMETRY: MW=7287.96; MW_ERR=0.37; METHOD=Electrospray.
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC       ALPHA-TOXIN SUBFAMILY. ALPHA-LIKE TOXIN SUBCLASS.
DR   Pfam; PF00537; toxin_3; 1.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor.
FT   DISULFID     12     63       BY SIMILARITY.
FT   DISULFID     16     36       BY SIMILARITY.
FT   DISULFID     22     46       BY SIMILARITY.
FT   DISULFID     26     48       BY SIMILARITY.
SQ   SEQUENCE   65 AA;  7296 MW;  4305EB32436A4B90 CRC64;

Alignment Scores:
Pred. No.:              0.0204        Length:        65
Score:                  78.50         Matches:       16
Percent Similarity:     51.22%        Conservative:  5
Best Local Similarity:  39.02%        Mismatches:    15
Query Match:            29.40%        Indels:        5
DB:                     1             Gaps:          3

US-09-673-274B-2_COPY_1_132 (1-132) x SCX4_BUTOM (1-65)

Qy         16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
              :::||||||:::                  |||   ||:|||      |||   :::
Db         11 AsnCysValTyrGluCysAlaLys---AsnSerTyrCysAsnAspLeuCysThrLysAsn 29

Qy         76 GGTTACAAGGGTGGTCACTGC---------GGTTCCTTCGCTAACGTGAACTGCTGGTGC 126
              |||   |||   |||:::|||         |||   :::   ||:      |||||||:|
Db         30 GlyAlaLysSerGlyTyrCysGlnTrpLeuGlyLysTyrGlyAsnAla---CysTrpCys 48

Qy        127 GAG 129
              |:|
Db         49 Glu 49

RESULT 9
SCX3_ANDAU
ID   SCX3_ANDAU     STANDARD;      PRT;    84 AA.
AC   P01480;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Neurotoxin III precursor (AaH III) (AaH3).
OS   Androctonus australis hector (Sahara scorpion).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Androctonus.
OX   NCBI_TaxID=70175;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90037062; PubMed=2808423;
RA   Bougis P.E., Rochat H., Smith L.A.;
RT   "Precursors of Androctonus australis scorpion neurotoxins. Structures
RT   of precursors, processing outcomes, and expression of a functional
RT   recombinant toxin II.";
RL   J. Biol. Chem. 264:19259-19265(1989).
RN   [2]
RP   SEQUENCE OF 20-83.
RX   MEDLINE=79148414; PubMed=428402;
RA   Kopeyan C., Martinez G., Rochat H.;
RT   "Amino acid sequence of neurotoxin III of the scorpion Androctonus
RT   australis Hector.";
RL   Eur. J. Biochem. 94:609-615(1979).
RN   [3]
RP   STRUCTURE BY NMR.
RX   MEDLINE=91122921; PubMed=2279844;
RA   Laplante S.R., Mikou A., Robin M., Guittet E., Delsuc M.A.,
RA   Charpentier I., Lallemand J.-Y.;
RT   "Rapid determination and NMR assignments of antiparallel sheets and
RT   helices of a scorpion and a cobra toxin.";
RL   Int. J. Pept. Protein Res. 36:227-230(1990).
RN   [4]
RP   STRUCTURE BY NMR.
RX   MEDLINE=93044377; PubMed=1422146;
RA   Mikou A., Laplante S.R., Guittet E., Lallemand J.-Y.,
RA   Martin-Eauclaire M.-F., Rochat H.;
RT   "Toxin III of the scorpion Androctonus australis hector: proton
RT   nuclear magnetic resonance assignments and secondary structure.";
RL   J. Biomol. NMR 2:57-70(1992).
CC   -!- FUNCTION: Binds to sodium channels and inhibits the inactivation
CC       of the activated channels, thereby blocking neuronal transmission.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC       ALPHA-TOXIN SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M27703; AAA29948.1; -.
DR   PIR; C34444; NTSR3A.
DR   HSSP; P56678; 1BMR.
DR   InterPro; IPR003614; Knot1.
DR   InterPro; IPR002061; Scorpion_toxinL.
DR   Pfam; PF00537; toxin_3; 1.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor;
KW   Signal.
FT   SIGNAL        1     19
FT   CHAIN        20     83       NEUROTOXIN III.
FT   PROPEP       84     84
FT   DISULFID     31     81       BY SIMILARITY.
FT   DISULFID     35     53       BY SIMILARITY.
FT   DISULFID     39     63       BY SIMILARITY.
FT   DISULFID     43     65       BY SIMILARITY.
FT   CONFLICT     27     27       D -> N (IN REF. 2).
SQ   SEQUENCE   84 AA;  9068 MW;  ADB0AD462153E3C0 CRC64;

Alignment Scores:
Pred. No.:              0.0207        Length:        84
Score:                  78.50         Matches:       15
Percent Similarity:     53.85%        Conservative:  6
Best Local Similarity:  38.46%        Mismatches:    13
Query Match:            29.40%        Indels:        5
DB:                     1             Gaps:          2

US-09-673-274B-2_COPY_1_132 (1-132) x SCX3_ANDAU (1-84)

Qy         16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
              :::||||||:::      |||               |||:::|||   ||||:|:::
Db         30 AsnCysValTyrHisCysVal---------ProProCysAspGlyLeuCysLysLysAsn 46

Qy         76 GGTTACAAGGGTGGTCACTGCGGTTCCTTC------GCTAACGTGAACTGCTGGTGC 126
              |||   |||   :||   ||||||            :::   :::   ||||||:::
Db         47 GlyAlaLysSerGlySerCysGlyPheLeuIleProSerGlyLeuAlaCysTrpCys 65

RESULT 10
SCL6_LEIQH
```

```
ID   SCL6_LEIQH     STANDARD;      PRT;    64 AA.
AC   P59356;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Alpha like neurotoxin Lqh VI (Lqh6).
OS   Leiurus quinquestriatus hebraeus (Yellow scorpion).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Leiurus.
OX   NCBI_TaxID=6884;
RN   [1]
RP   SEQUENCE, PHARMACOLOGICAL CHARACTERIZATION, AND BIOASSAY.
RC   TISSUE=Venom;
RX   MEDLINE=22168983; PubMed=12180969;
RA   Hamon A., Gilles N., Sautiere P., Martinage A., Kopeyan C., Ulens C.,
RA   Tytgat J., Lancelin J.-M., Gordon D.;
RT   "Characterization of scorpion alpha-like toxin group using two new
RT   toxins from the scorpion Leiurus quinquestriatus hebraeus.";
RL   Eur. J. Biochem. 269:3920-3933(2002).
CC   -!- FUNCTION: Binds to sodium channels and inhibits the inactivation
CC       of the activated channels, thereby blocking neuronal transmission.
CC       This toxin is highly toxic to insects and mice, and inhibits the
CC       binding of alpha-toxin to cockroach neuronal membranes.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC   -!- MISCELLANEOUS: LD(50) is 34.3 nmol/kg to cockroaches (Blatella
CC       germanica).
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC       ALPHA-TOXIN SUBFAMILY. ALPHA-LIKE TOXIN SUBCLASS.
DR   Pfam; PF00537; toxin_3; 1.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor;
KW   Amidation.
FT   DISULFID     12     62       BY SIMILARITY.
FT   DISULFID     16     34       BY SIMILARITY.
FT   DISULFID     20     44       BY SIMILARITY.
FT   DISULFID     24     46       BY SIMILARITY.
FT   MOD_RES      64     64       AMIDATION.
SQ   SEQUENCE   64 AA;  6803 MW;  49A54B0B6C75392E CRC64;

Alignment Scores:
Pred. No.:              0.0276        Length:          64
Score:                  77.50         Matches:         15
Percent Similarity:     51.28%        Conservative:    5
Best Local Similarity:  38.46%        Mismatches:      14
Query Match:            29.03%        Indels:          5
DB:                     1             Gaps:            2

US-09-673-274B-2_COPY_1_132 (1-132) x SCL6_LEIQH (1-64)

Qy         16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
              :::||||||:::      :::            ||||||:::      ||||||
Db         11 AsnCysValTyrHisCysIle---------ProAspCysAspThrLeuCysLysAspAsn 27

Qy         76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCT------AACGTGAACTGCTGGTGC 126
              ||.      |||||||||||||||                  :::   |||||||||
Db         28 GlyGlyThrGlyGlyHisCysGlyPheLysLeuGlyHisGlyIleAlaCysTrpCys 46

RESULT 11
SCX4_LEIQU
ID   SCX4_LEIQU     STANDARD;      PRT;    65 AA.
AC   P01489;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Neurotoxin IV (Lqq IV) (Lqq4).
OS   Leiurus quinquestriatus quinquestriatus (Egyptian scorpion).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Leiurus.
OX   NCBI_TaxID=6885;
RN   [1]
RP   SEQUENCE.
RA   Kopeyan C., Martinez G., Rochat H.;
RT   "Primary structure of toxin IV of Leiurus quinquestriatus
RT   quinquestriatus.";
RL   FEBS Lett. 181:211-217(1985).
CC   -!- FUNCTION: Binds to sodium channels and inhibits the inactivation
CC       of the activated channels, thereby blocking neuronal transmission.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC       ALPHA-TOXIN SUBFAMILY.
DR   PIR; A01749; NTSR4L.
DR   HSSP; P45697; 1SN1.
DR   InterPro; IPR003614; Knot1.
DR   InterPro; IPR001219; Neurotoxin.
DR   InterPro; IPR002061; Scorpion_toxinL.
DR   Pfam; PF00537; toxin_3; 1.
DR   PRINTS; PR00284; TOXIN.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor;
KW   Amidation.
FT   DISULFID     13     64       BY SIMILARITY.
FT   DISULFID     17     37       BY SIMILARITY.
FT   DISULFID     23     47       BY SIMILARITY.
FT   DISULFID     27     49       BY SIMILARITY.
FT   MOD_RES      65     65       AMIDATION.
SQ   SEQUENCE   65 AA;  7202 MW;  882102EC1C2B89F6 CRC64;

Alignment Scores:
Pred. No.:              0.0276        Length:          65
Score:                  77.50         Matches:         16
Percent Similarity:     57.14%        Conservative:    8
Best Local Similarity:  38.10%        Mismatches:      9
Query Match:            29.03%        Indels:          9
DB:                     1             Gaps:            4

US-09-673-274B-2_COPY_1_132 (1-132) x SCX4_LEIQU (1-65)

Qy         16 TCCTGCGTGTGG------GGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAG 69
              :::||||||:::      |||:::   :::|||         ||||||   ||||||
Db         12 AsnCysValTyrThrCysGlySerAsnSerTyr---------CysAsnThrGluCysThr 28

Qy         70 AGGAGGGGTTACAAGGGTGGTCACTGC---------GGTTCCTTCGCTAACGTGAACTGC 120
              :::   :||   :::   |||:::|||         |||   :::   |||      |||
Db         29 LysAsnGlyAlaGluSerGlyTyrCysGlnTrpLeuGlyLysTyrGlyAsnAla---Cys 47

Qy        121 TGGTGC 126
              ||||||
Db         48 TrpCys 49

RESULT 12
SCAT_MESMA
ID   SCAT_MESMA     STANDARD;      PRT;    85 AA.
AC   Q9UAC8;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Neurotoxin AS-1 precursor.
OS   Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Mesobuthus.
OX   NCBI_TaxID=34649;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Venom gland;
RA   Lan Z.D., Dai L., Chi C.-W.;
RT   "The gene cloning and sequencing of AS and AS-1, two novel neurotoxins
RT   from the scorpion Buthus martensi Karsch.";
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Ryanodine receptor agonist in skeletal muscle.
CC   -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC   -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
```

```
CC       ALPHA-TOXIN SUBFAMILY.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; AF079061; AAD47375.1; -.
DR   HSSP; P01492; 1VNB.
DR   InterPro; IPR003614; Knot1.
DR   InterPro; IPR002061; Scorpion_toxinL.
DR   Pfam; PF00537; toxin_3; 1.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Toxin; Neurotoxin; Signal.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20     85       NEUROTOXIN AS-1.
FT   DISULFID     31     81       BY SIMILARITY.
FT   DISULFID     35     56       BY SIMILARITY.
FT   DISULFID     42     63       BY SIMILARITY.
FT   DISULFID     46     65       BY SIMILARITY.
SQ   SEQUENCE   85 AA;  9802 MW;  E2274838439E2B95 CRC64;

Alignment Scores:
Pred. No.:              0.0327        Length:         85
Score:                  77.00         Matches:        17
Percent Similarity:     53.49%        Conservative:   6
Best Local Similarity:  39.53%        Mismatches:     18
Query Match:            28.84%        Indels:         2
DB:                     1             Gaps:           2

US-09-673-274B-2_COPY_1_132 (1-132) x SCAT_MESMA (1-85)

Qy            1 GATAAGCTTATCGGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
                :::|||      |||      :::|||      :::|||      '!|   ||||||
Db           26 AsnLysTyrThrGlyCysLysIleTrpCysValIleAsnAsnGluSer---CysAsnSer 44

Qy           61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
                |||||||||   |||            |||:::|||      |||      :::   |||
Db           45 GluCysLysLeuArgArgGlyAsnTyrGlyTyrCys---TyrPheTrpLysLeuAlaCys 63

Qy          121 TGGTGCGAG 129
                :::||||||
Db           64 TyrCysGlu 66

RESULT 13
SIX2_LEIQU
ID   SIX2_LEIQU     STANDARD;      PRT;    61 AA.
AC   P19855;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Insect toxin 2 (Insect toxin LqqIT2).
OS   Leiurus quinquestriatus quinquestriatus (Egyptian scorpion).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Leiurus.
OX   NCBI_TaxID=6885;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=90184494; PubMed=2311768;
RA   Kopeyan C., Mansuelle P., Sampieri F., Brando T., Bahraoui E.M.,
RA   Rochat H., Granier C.;
RT   "Primary structure of scorpion anti-insect toxins isolated from the
RT   venom of Leiurus quinquestriatus quinquestriatus.";
RL   FEBS Lett. 261:423-426(1990).
RN   [2]
RP   SEQUENCE.
RX   MEDLINE=93160530; PubMed=8431601;
RA   Zlotkin E., Gurevitz M., Fowler E., Adams M.E.;
RT   "Depressant insect selective neurotoxins from scorpion venom:
RT   chemistry, action, and gene cloning.";
RL   Arch. Insect Biochem. Physiol. 22:55-73(1993).
RN   [3]
RP   ACTION.
RA   Zlotkin E., Fowler E., Eitan M., Moyer M., Adams M.E.;
RT   "On the chemistry and action of the depressant insect toxins.";
RL   Toxicon 28:170-170(1990).
CC   -!- FUNCTION: Binds to sodium channels and inhibits the inactivation
CC       of the activated channels, thereby blocking neuronal transmission.
CC       This toxin induces a slow, depressant, flaccid paralysis on fly
CC       larvae. It is active only on insects.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC       ALPHA-TOXIN SUBFAMILY.
DR   PIR; B34123; B34123.
DR   HSSP; P01494; 2SN3.
DR   InterPro; IPR003614; Knot1.
DR   InterPro; IPR002061; Scorpion_toxinL.
DR   Pfam; PF00537; toxin_3; 1.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor.
FT   DISULFID     10     60       BY SIMILARITY.
FT   DISULFID     14     35       BY SIMILARITY.
FT   DISULFID     21     42       BY SIMILARITY.
FT   DISULFID     25     44       BY SIMILARITY.
SQ   SEQUENCE   61 AA;  6860 MW;  80F01771C7DCFAF6 CRC64;

Alignment Scores:
Pred. No.:              0.0372        Length:         61
Score:                  76.50         Matches:        15
Percent Similarity:     55.26%        Conservative:   6
Best Local Similarity:  39.47%        Mismatches:     12
Query Match:            28.65%        Indels:         5
DB:                     1             Gaps:           2

US-09-673-274B-2_COPY_1_132 (1-132) x SIX2_LEIQU (1-61)

Qy           16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
                ||||||::::::|||                     .|||||   ||||.:':|
Db           13 SerCysLeuPheGly------------AsnGluGlyCysAsnLysGluCysLysSerTyr 28

Qy           76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
                |||         |||:::|||   ::::::      :::   ||:||||||' |
Db           29 GlyGlySerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 45

RESULT 14
SCXE_BUTEU
ID   SCXE_BUTEU     STANDARD;      PRT;    66 AA.
AC   P09982;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-MAR-1989 (Rel. 10, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Neurotoxin M14.
OS   Buthus eupeus (Lesser Asian scorpion) (Mesobuthus eupeus).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Mesobuthus.
OX   NCBI_TaxID=34648;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Venom;
RX   MEDLINE=85046632; PubMed=6497916;
RA   Volkova T.M., Garsia A.F., Telezhinskaya I.N., Potapenko N.A.,
RA   Grishin E.V.;
RT   "Amino acid sequence of 2 neurotoxins from the scorpion Buthus eupeus
RT   venom.";
RL   Bioorg. Khim. 10:979-982(1984).
CC   -!- FUNCTION: Binds to sodium channels and inhibits the inactivation
CC       of the activated channels, thereby blocking neuronal transmission.
CC       Has paralytic activity in mice.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom gland.
```

```
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC       ALPHA-TOXIN SUBFAMILY.
DR   PIR; JT0020; NTSR4E.
DR   HSSP; P45697; 1SN1.
DR   InterPro; IPR003614; Knot1.
DR   InterPro; IPR001219; Neurotoxin.
DR   InterPro; IPR002061; Scorpion_toxinL.
DR   Pfam; PF00537; toxin_3; 1.
DR   PRINTS; PR00284; TOXIN.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor.
FT   DISULFID     12     65       BY SIMILARITY.
FT   DISULFID     16     36       BY SIMILARITY.
FT   DISULFID     22     46       BY SIMILARITY.
FT   DISULFID     26     48       BY SIMILARITY.
SQ   SEQUENCE   66 AA;  7453 MW;  B5E0DB42237EF2E8 CRC64;

Alignment Scores:
Pred. No.:                 0.0435         Length:         66
Score:                     76.00          Matches:        13
Percent Similarity:        58.62%         Conservative:   4
Best Local Similarity:     44.83%         Mismatches:     8
Query Match:               28.46%         Indels:         4
DB:                        1              Gaps:           2

US-09-673-274B-2_COPY_1_132 (1-132) x SCXE_BUTEU (1-66)

Qy         52 TGCAACGGTGAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGC---------GGTTCC 102
              ||::::    :|:||||||:::  : |      ||.   ::: ||         |||
Db         22 CysAspSerGluCysLysLysAsnGlyAlaAspGlySerTyrCysGlnTrpLeuGlyArg 41

Qy        103 TTCGCTAACGTGAACTGCTGGTGCGAG 129
              |||   |||      |||:|||||:::
Db         42 PheGlyAsnAla---CysTrpCysLys 49

RESULT 15
SCL3_LEIQH
ID   SCL3_LEIQH     STANDARD;      PRT;    67 AA.
AC   P56678;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Alpha-like neurotoxin Lqh III (Lqh3).
OS   Leiurus quinquestriatus hebraeus (Yellow scorpion).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Leiurus.
OX   NCBI_TaxID=6884;
RN   [1]
RP   SEQUENCE, PHARMACOLOGICAL CHARACTERIZATION, AND BIOASSAY.
RC   TISSUE=Venom;
RX   MEDLINE=98353091; PubMed=9690781;
RA   Sautiere P., Cestele S., Kopeyan C., Martinage A., Drobecq H.,
RA   Doljansky Y., Gordon D.;
RT   "New toxins acting on sodium channels from the scorpion Leiurus
RT   quinquestriatus hebraeus suggest a clue to mammalian vs insect
RT   selectivity.";
RL   Toxicon 36:1141-1154(1998).
RN   [2]
RP   STRUCTURE BY NMR.
RC   TISSUE=Venom;
RX   MEDLINE=99141414; PubMed=9917409;
RA   Krimm I., Gilles N., Sautiere P., Stankiewicz M., Pelhate M.,
RA   Gordon D., Lancelin J.-M.;
RT   "NMR structures and activity of a novel alpha-like toxin from the
RT   scorpion Leiurus quinquestriatus hebraeus.";
RL   J. Mol. Biol. 285:1749-1763(1999).
RN   [3]
RP   STRUCTURE BY NMR.
RC   TISSUE=Venom;
RA   Krimm I., Trivelli X., Lancelin J.-M.;
RT   "A cis trans isomerism of a non-prolyl peptide bond in LQH III alpha-
RT   like scorpion toxin revealed by solution NMR.";
```

```
RL   Submitted (JUL-2000) to the PDB data bank.
CC   -!- FUNCTION: Binds to sodium channels and inhibits the inactivation
CC       of the activated channels, thereby blocking neuronal transmission.
CC       Active on both insects and mammals but competes for alpha-toxins
CC       binding only on cockroach sodium channels.
CC   -!- SUBUNIT: Monomer.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC   -!- MISCELLANEOUS: LD(50) is 50 mg/kg by intracerebroventricular
CC       injection into mice.
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC       ALPHA-TOXIN SUBFAMILY. ALPHA-LIKE TOXIN SUBCLASS.
DR   PDB; 1BMR; 16-FEB-99.
DR   PDB; 1FH3; 08-NOV-00.
DR   InterPro; IPR003614; Knot1.
DR   InterPro; IPR002061; Scorpion_toxinL.
DR   Pfam; PF00537; toxin_3; 1.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor;
KW   Amidation; 3D-structure.
FT   DISULFID     12     65
FT   DISULFID     16     37
FT   DISULFID     23     47
FT   DISULFID     27     49
FT   MOD_RES      67     67       AMIDATION.
FT   STRAND        2      6
FT   TURN          9     11
FT   HELIX        21     30
FT   TURN         31     31
FT   STRAND       34     40
FT   TURN         41     43
FT   STRAND       44     52
FT   TURN         54     55
SQ   SEQUENCE   67 AA;  7057 MW;  19FE8EF96154328F CRC64;

Alignment Scores:
Pred. No.:                 0.0436         Length:         67
Score:                     76.00          Matches:        14
Percent Similarity:        51.28%         Conservative:   6
Best Local Similarity:     35.90%         Mismatches:     17
Query Match:               28.46%         Indels:         2
DB:                        1              Gaps:           1

US-09-673-274B-2_COPY_1_132 (1-132) x SCL3_LEIQH (1-67)

Qy         16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
              :::||||||:::               :::|||   |||:::      ||||||   :::
Db         11 AsnCysValTyrHisCysPheProGlySerSerGlyCysAspThrLeuCysLysGluLys 30

Qy         76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCT------AACGTGAACTGCTGGTGC 126
              ||:         ||||||.||||:                  :::   |||||||||
Db         31 GlyGlyThrSerGlyHisCysGlyPheLysValGlyHisGlyLeuAlaCysTrpCys 49

Search completed: October 16, 2003, 17:07:24
Job time : 9.77076 secs
```

```
                       GenCore version 5.1.6
               Copyright (c) 1993 - 2003  Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on:         October 16, 2003, 17:06:33 ; Search time 40.7841 Seconds
                                             (without alignments)
                                             1027.456 Million cell updates/sec

Title:          US-09-673-274B-2_COPY_1_132
Perfect score:  267
Sequence:       1 gataagcttatcggttcctg..........tgaactgctggtgcgagact 132

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters:      2215726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09673274/runat_16102003_170626_2523/app_query.fasta_1.654
-DB=A_Geneseq_19Jun03 -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09673274_@CGN_1_1_195_@runat_16102003_170626_2523 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      A_Geneseq_19Jun03:*
                1:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
                2:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
                3:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
                4:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
                5:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
                6:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
                7:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
                8:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
                9:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
                10:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
                11:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
                12:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
                13:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
                14:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
                15:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
                16:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
                17:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
                18:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
                19:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
                20:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
                21:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
                22:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
                23:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
                24:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

    Pred. No. is the number of results predicted by chance to have a
    score greater than or equal to the score of the result being printed,
    and is derived by analysis of the total score distribution.
```

```
                                      SUMMARIES
                       %
Result               Query
  No.     Score      Match Length DB  ID                             Description
-------------------------------------------------------------------------------------
     1      267      100.0     44  21  AAY54375                      Amino acid sequenc
     2      267      100.0     44  23  ABB76860                      Antifungal/antibac
     3      267      100.0     44  23  ABB76868                      Antifungal/antibac
     4      267      100.0     49  21  AAY54374                      Amino acid sequenc
     5      266       99.6     44  23  ABB76895                      Antifungal/antibac
     6      264       98.9     44  23  ABB76863                      Antifungal/antibac
     7      264       98.9     44  23  ABB76864                      Antifungal/antibac
     8      264       98.9     44  23  ABB76872                      Antifungal/antibac
     9      264       98.9     44  23  ABB76896                      Antifungal/antibac
    10      262       98.1     44  23  ABB76862                      Antifungal/antibac
    11      262       98.1     44  23  ABB76865                      Antifungal/antibac
    12      262       98.1     44  23  ABB76867                      Antifungal/antibac
    13      262       98.1     44  23  ABB76869                      Antifungal/antibac
    14      262       98.1     44  23  ABB76871                      Antifungal/antibac
    15      262       98.1     74  21  AAY54376                      A fusion peptide o
    16      261       97.8     44  23  ABB76866                      Antifungal/antibac
    17      261       97.8     44  23  ABB76870                      Antifungal/antibac
    18      261       97.8     44  23  ABB76894                      Antifungal/antibac
    19      259       97.0     44  23  ABB76873                      Antifungal/antibac
    20      258       96.6     44  23  ABB76898                      Antifungal/antibac
    21      256       95.9     44  23  ABB76861                      Antifungal/antibac
    22      256       95.9     44  23  ABB76897                      Antifungal/antibac
    23      255       95.5     44  23  ABB76882                      Antifungal/antibac
    24      254       95.1     44  23  ABB76874                      Antifungal/antibac
    25      254       95.1     44  23  ABB76875                      Antifungal/antibac
    26      253       94.8     44  23  ABB76879                      Antifungal/antibac
    27      253       94.8     44  23  ABB76885                      Antifungal/antibac
    28      251       94.0     44  23  ABB76876                      Antifungal/antibac
    29      251       94.0     44  23  ABB76877                      Antifungal/antibac
    30      251       94.0     44  23  ABB76881                      Antifungal/antibac
    31      251       94.0     44  23  ABB76884                      Antifungal/antibac
    32      251       94.0     44  23  ABB76886                      Antifungal/antibac
    33      250       93.6     44  23  ABB76880                      Antifungal/antibac
    34      250       93.6     44  23  ABB76883                      Antifungal/antibac
    35      248       92.9     44  23  ABB76887                      Antifungal/antibac
    36      248       92.9     44  23  ABB76888                      Antifungal/antibac
    37      246       92.1     44  23  ABB76878                      Antifungal/antibac
    38      246       92.1     44  23  ABB76889                      Antifungal/antibac
    39      246       92.1     44  23  ABB76890                      Antifungal/antibac
    40      243       91.0     44  23  ABB76892                      Antifungal/antibac
    41      243       91.0     44  23  ABB76899                      Antifungal/antibac
    42      241       90.3     44  23  ABB76891                      Antifungal/antibac
    43      238       89.1     44  23  ABB76893                      Antifungal/antibac
    44      232       86.9     44  24  ABP55988                      Agrotis ipsilon Fu
    45      232       86.9     73  24  ABP55987                      Agrotis ipsilon Fu


                                     ALIGNMENTS

RESULT 1
AAY54375
ID   AAY54375 standard; Peptide; 44 AA.
XX
AC   AAY54375;
XX
DT   06-APR-2000  (first entry)
XX
DE   Amino acid sequence of a heliomycin peptide.
XX
KW   Heliomycin; polycysteine peptide; antifungal; antibacterial;
KW   fungal infection; plant protection; animal protection;
KW   transgenic plant; fungi; bacteria; Cercospora beticola;
KW   Cladosporium herbarum; Fusarium culmorum; F. oxysporum;
KW   Phytophthora cinnamoni.
XX
CS   Heliothis virescens.
XX
```

```
PN   FR2777568 A1.
XX
PD   22 OCT 1999.
XX
PF   15 APR 1998;    98FR-0004933.
XX
PR   15-APR 1998;    98FR-0004933.
XX
PA   (RHON ) RHONE-POULENC AGROCHIMIE.
XX
DR   WPI; 2000-108532/10.
DR   N PSDB; AAZ45692.
XX
PT   New polycysteine peptides, designated heliomycin, with antifungal or
PT   antibacterial activity, useful in medicine or plant protection  -
XX
PS   Claim 7; Page 31-32; 46pp; French.
XX
CC   The present sequence represents the heliomycin peptide. Heliomycin is a
CC   polycysteine peptide with antifungal and antibacterial activity,
CC   isolated from the lepidopteron Heliothis virescens. Heliomycin peptides
CC   contain the sequence given in AAY54378. The heliomycin peptides are
CC   used as pharmaceuticals for treating or preventing fungal infections
CC   in humans and animals. They may also be used as antifungal agents for
CC   plant protection. Nucleic acid encoding the heliomycin peptides is
CC   used to generate transgenic plants that are resistant to some fungi
CC   and bacteria (specifically Cercospora beticola, Cladosporium herbarum,
CC   Fusarium culmorum, F. oxysporum and Phytophthora cinnamoni), and for
CC   recombinant production of the peptides.
XX
SQ   Sequence   44 AA;

Alignment Scores:
Pred. No.:               4.66e-27       Length:         44
Score:                   267.00         Matches:        44
Percent Similarity:      100.00%        Conservative:   0
Best Local Similarity:   100.00%        Mismatches:     0
Query Match:             100.00%        Indels:         0
DB:                      21             Gaps:           0

US-09-673-274B-2_COPY_1_132 (1-132) x AAY54375 (1-44)

Qy          1 GATAAGCTTATCGGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AspLysLeuIleGlySerCysValTrpGlyAlaValAsnTyrThrSerAspCysAsnGly 20
                       6                     15
Qy         61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 GluCysLysArgArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnValAsnCys 40
              3                        9                        7
Qy        121 TGGTGCGAGACT 132
              ||||||||||||
Db         41 TrpCysGluThr 44
                       2
RESULT 2
ABB76860
ID   ABB76860 standard; Protein; 44 AA.
XX
AC   ABB76860;
XX
DT   03-JUL-2002  (first entry)
XX
DE   Antifungal/antibacterial peptide Helio.
XX
KW   Antifungal; antibacterial; fungicide; heliomicine; human medicine;
KW   veterinary medicine; plant protection; bacterial infection;
KW   fungal infection.
XX
OS   Heliothis virescens.
XX
PN   WO200206324-A2.
XX
PD   24-JAN-2002.
XX
PF   05-JUL-2001; 2001WO-FR02164.
XX
PR   13-JUL-2000; 2000FR-0009248.
PR   19-SEP-2000; 2000FR-0011949.
XX
PA   (ENTO-) ENTOMED SA.
XX
PI   Dimarcq J,  Legrain M,  Menin L;
XX
DR   WPI; 2002-179779/23.
XX
PT   New peptide derivatives of heliomicine, useful as antibacterial and
PT   antifungal agents in human or veterinary medicine and agriculture  -
XX
PS   Claim 15; Page 57; 100pp; French.
XX
CC   The present invention relates to peptide derivatives of heliomicine with
CC   antifungal and/or antibacterial activity. The present sequence is one
CC   such peptide. The peptides are useful in human and veterinary medicine
CC   and in plant protection, for control of bacterial and fungal infections.
CC   An antifungal and/or antibacterial composition containing at least one
CC   peptide is also claimed.
XX
SQ   Sequence   44 AA;

Alignment Scores:
Pred. No.:               4.66e-27       Length:         44
Score:                   267.00         Matches:        44
Percent Similarity:      100.00%        Conservative:   0
Best Local Similarity:   100.00%        Mismatches:     0
Query Match:             100.00%        Indels:         0
DB:                      23             Gaps:           0

US-09-673-274B-2_COPY_1_132 (1-132) x ABB76860 (1-44)

Qy          1 GATAAGCTTATCGGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AspLysLeuIleGlySerCysValTrpGlyAlaValAsnTyrThrSerAspCysAsnGly 20

Qy         61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 GluCysLysArgArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnValAsnCys 40

Qy        121 TGGTGCGAGACT 132
              ||||||||||||
Db         41 TrpCysGluThr 44

RESULT 3
ABB76868
ID   ABB76868 standard; Protein; 44 AA.
XX
AC   ABB76868;
XX
DT   03-JUL-2002  (first entry)
XX
DE   Antifungal/antibacterial peptide pEM23.
XX
KW   Antifungal; antibacterial; fungicide; heliomicine; human medicine;
KW   veterinary medicine; plant protection; bacterial infection;
KW   fungal infection.
XX
OS   Synthetic.
XX
PN   WO200206324-A2.
XX
PD   24-JAN-2002.
XX
PF   05-JUL-2001; 2001WO-FR02164.
XX
PR   13-JUL-2000; 2000FR-0009248.
PR   19-SEP-2000; 2000FR-0011949.
XX
```

```
PA   (ENTO-) ENTOMED SA.
XX
PI   Dimarcq J,  Legrain M,  Menin L;
XX
DR   WPI; 2002-179779/23.
XX
PT   New peptide derivatives of heliomicine, useful as antibacterial and
PT   antifungal agents in human or veterinary medicine and agriculture  -
XX
PS   Claim 15; Page 57; 100pp; French.
XX
CC   The present invention relates to peptide derivatives of heliomicine with
CC   antifungal and/or antibacterial activity. The present sequence is one
CC   such peptide. The peptides are useful in human and veterinary medicine
CC   and in plant protection, for control of bacterial and fungal infections.
CC   An antifungal and/or antibacterial composition containing at least one
CC   peptide is also claimed.
XX
SQ   Sequence   44 AA;

Alignment Scores:
Pred. No.:              4.66e-27      Length:         44
Score:                  267.00        Matches:        44
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            100.00%       Indels:         0
DB:                     23            Gaps:           0

US-09-673-274B-2_COPY_1_132 (1-132) x ABB76868 (1-44)

Qy          1 GATAAGCTTATCGGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AspLysLeuIleGlySerCysValTrpGlyAlaValAsnTyrThrSerAspCysAsnGly 20

Qy         61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 GluCysLysArgArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnValAsnCys 40

Qy        121 TGGTGCGAGACT 132
              ||||||||||||
Db         41 TrpCysGluThr 44

RESULT 4
AAY54374
ID   AAY54374 standard; Peptide; 49 AA.
XX
AC   AAY54374;
XX
DT   06-APR-2000  (first entry)
XX
DE   Amino acid sequence of a fusion peptide of MF-alpha1/heliomycin.
XX
KW   Mat alpha1 factor; MF-alpha1; heliomycin; polycysteine peptide;
KW   antifungal; antibacterial; fungal infection; plant protection;
KW   animal protection; transgenic plant; fungi; bacteria;
KW   Cercospora beticola; Cladosporium herbarum; Fusarium culmorum;
KW   F. oxysporum; Phytophthora cinnamoni.
XX
OS   Chimeric - Saccharomyces cerevisiae.
OS   Chimeric - Heliothis virescens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..5
FT                   /note= "residues derived from the yeast Mat-alpha1
FT                          (MF-alpha1) factor"
FT   Peptide         6..49
FT                   /note= "heliomycin residues"
XX
PN   FR2777568-A1.
XX
PD   22-OCT-1999.
XX
PF   15-APR-1998;   98FR-0004933.
```

```
XX
PR   15-APR-1998;   98FR-0004933.
XX
PA   (RHON ) RHONE-POULENC AGROCHIMIE.
XX
DR   WPI; 2000-108532/10.
DR   N-PSDB; AAZ45691.
XX
PT   New polycysteine peptides, designated heliomycin, with antifungal or
PT   antibacterial activity, useful in medicine or plant protection  -
XX
PS   Claim 13; Page 31; 46pp; French.
XX
CC   The present sequence represents a fusion peptide of 5 amino acids
CC   derived from the Mat alpha1 factor (MF-alpha1) and heliomycin. The
CC   attachment of the MF-alpha1 sequence to the N-terminal aids secretion
CC   of the heliomycin peptide. Heliomycin is a polycysteine peptide with
CC   antifungal and antibacterial activity, isolated from the lepidopteron
CC   Heliothis virescens. Heliomycin peptides contain the sequence given in
CC   AAY54378. The heliomycin peptides are used as pharmaceuticals for
CC   treating or preventing fungal infections in humans and animals. They
CC   may also be used as antifungal agents for plant protection. Nucleic
CC   acid encoding the heliomycin peptides is used to generate transgenic
CC   plants that are resistant to some fungi and bacteria (specifically
CC   Cercospora beticola, Cladosporium herbarum, Fusarium culmorum,
CC   F. oxysporum and Phytophthora cinnamoni), and for recombinant production
CC   of the peptides.
XX
SQ   Sequence   49 AA;

Alignment Scores:
Pred. No.:              4.73e-27      Length:         49
Score:                  267.00        Matches:        44
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            100.00%       Indels:         0
DB:                     21            Gaps:           0

US-09-673-274B-2_COPY_1_132 (1-132) x AAY54374 (1-49)

Qy          1 GATAAGCTTATCGGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6 AspLysLeuIleGlySerCysValTrpGlyAlaValAsnTyrThrSerAspCysAsnGly 25

Qy         61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 GluCysLysArgArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnValAsnCys 45

Qy        121 TGGTGCGAGACT 132
              ||||||||||||
Db         46 TrpCysGluThr 49

RESULT 5
ABB76895
ID   ABB76895 standard; Protein; 44 AA.
XX
AC   ABB76895;
XX
DT   03-JUL-2002  (first entry)
XX
DE   Antifungal/antibacterial peptide #2.
XX
KW   Antifungal; antibacterial; fungicide; heliomicine; human medicine;
KW   veterinary medicine; plant protection; bacterial infection;
KW   fungal infection.
XX
OS   Synthetic.
XX
PN   WO200206324-A2.
XX
PD   24-JAN-2002.
XX
PF   05-JUL-2001; 2001WO-FR02164.
```

```
XX
PR   13-JUL-2000; 2000FR-0009248.
PR   19 SEP 2000; 2000FR-0011949.
XX
PA   (ENTO-) ENTOMED SA.
XX
PI   Dimarcq J,  Legrain M,  Menin L;
XX
DR   WPI; 2002-179779/23.
XX
PT   New peptide derivatives of heliomicine, useful as antibacterial and
PT   antifungal agents in human or veterinary medicine and agriculture  -
XX
PS   Disclosure; Page 86; 100pp; French.
XX
CC   The present invention relates to peptide derivatives of heliomicine with
CC   antifungal and/or antibacterial activity. The present sequence is one
CC   such peptide. The peptides are useful in human and veterinary medicine
CC   and in plant protection, for control of bacterial and fungal infections.
CC   An antifungal and/or antibacterial composition containing at least one
CC   peptide is also claimed.
XX
SQ   Sequence   44 AA;

Alignment Scores:
Pred. No.:                6.33e-27       Length:          44
Score:                    266.00         Matches:         43
Percent Similarity:       100.00%        Conservative:    1
Best Local Similarity:    97.73%         Mismatches:      0
Query Match:              99.63%         Indels:          0
DB:                       23             Gaps:            0

US-09-673-274B-2_COPY_1_132 (1-132) x ABB76895 (1-44)

Qy          1 GATAAGCTTATCGGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
              |:    |||||:|:||||||:|||||:| |||||:|||  |||:||||||:||||:|||:|
Db          1 AspLysLeuIleGlySerCysValTrpGlyAlaValAsnTyrThrSerAspCysAsnGly 20

Qy         61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
              ||::.|||||::||||||:|.|||||:.||||:.|||:|||||:||||:.||:::||||:|
Db         21 GluCysLysArgArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnIleAsnCys 40

Qy        121 TGGTGCGAGACT 132
              ::|||||:||||
Db         41 TrpCysGluThr 44

RESULT 6
ABB76863
ID   ABB76863 standard; Protein; 44 AA.
XX
AC   ABB76863;
XX
DT   03-JUL-2002  (first entry)
XX
DE   Antifungal/antibacterial peptide pEM38.
XX
KW   Antifungal; antibacterial; fungicide; heliomicine; human medicine;
KW   veterinary medicine; plant protection; bacterial infection;
KW   fungal infection.
XX
OS   Synthetic.
XX
PN   WO200206324-A2.
XX
PD   24 JAN 2002.
XX
PF   05-JUL 2001; 2001WO-FR02164.
XX
PR   13-JUL-2000; 2000FR-0009248.
PR   19-SEP-2000; 2000FR-0011949.
XX
PA   (ENTO-) ENTOMED SA.
XX
PI   Dimarcq J,  Legrain M,  Menin L;
XX
DR   WPI; 2002-179779/23.
XX
PT   New peptide derivatives of heliomicine, useful as antibacterial and
PT   antifungal agents in human or veterinary medicine and agriculture  -
XX
PS   Claim 15; Page 57; 100pp; French.
XX
CC   The present invention relates to peptide derivatives of heliomicine with
CC   antifungal and/or antibacterial activity. The present sequence is one
CC   such peptide. The peptides are useful in human and veterinary medicine
CC   and in plant protection, for control of bacterial and fungal infections.
CC   An antifungal and/or antibacterial composition containing at least one
CC   peptide is also claimed.
XX
SQ   Sequence   44 AA;

Alignment Scores:
Pred. No.:                1.17e-26       Length:          44
Score:                    264.00         Matches:         43
Percent Similarity:       100.00%        Conservative:    1
Best Local Similarity:    97.73%         Mismatches:      0
Query Match:              98.88%         Indels:          0
DB:                       23             Gaps:            0

US-09-673-274B-2_COPY_1_132 (1-132) x ABB76863 (1-44)

Qy          1 GATAAGCTTATCGGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
              ||||||||||||||||:::|||||||||||||||||||||||||||||||||||||:|||
Db          1 AspLysLeuIleGlyThrCysValTrpGlyAlaValAsnTyrThrSerAspCysAsnGly 20

Qy         61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
              ||||||||||||||||:|||:||||||||'||||||||||||||||||||||||||||||
Db         21 GluCysLysArgArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnValAsnCys 40

Qy        121 TGGTGCGAGACT 132
              ||||||||||||
Db         41 TrpCysGluThr 44

RESULT 7
ABB76864
ID   ABB76864 standard; Protein; 44 AA.
XX
AC   ABB76864;
XX
DT   03-JUL-2002  (first entry)
XX
DE   Antifungal/antibacterial peptide pEM43.
XX
KW   Antifungal; antibacterial; fungicide; heliomicine; human medicine;
KW   veterinary medicine; plant protection; bacterial infection;
KW   fungal infection.
XX
OS   Synthetic.
XX
PN   WO200206324-A2.
XX
PD   24-JAN-2002.
XX
PF   05-JUL-2001; 2001WO-FR02164.
XX
PR   13-JUL-2000; 2000FR-0009248.
PR   19-SEP-2000; 2000FR-0011949.
XX
PA   (ENTO-) ENTOMED SA.
XX
PI   Dimarcq J,  Legrain M,  Menin L;
XX
DR   WPI; 2002-179779/23.
XX
PT   New peptide derivatives of heliomicine, useful as antibacterial and
PT   antifungal agents in human or veterinary medicine and agriculture  -
```

```
X
S   Claim 15; Page 57; 100pp; French.
X
C   The present invention relates to peptide derivatives of heliomicine with
C   antifungal and/or antibacterial activity. The present sequence is one
C   such peptide. The peptides are useful in human and veterinary medicine
C   and in plant protection, for control of bacterial and fungal infections.
C   An antifungal and/or antibacterial composition containing at least one
C   peptide is also claimed.
X
Q   Sequence   44 AA;

lignment Scores:
red. No.:               1.17e-26       Length:        44
core:                   264.00         Matches:       43
ercent Similarity:      100.00%        Conservative:  1
est Local Similarity:   97.73%         Mismatches:    0
uery Match:             98.88%         Indels:        0
B:                      23             Gaps:          0

S-09-673-274B-2_COPY_1_132 (1-132) x ABB76864 (1-44)

y          1 GATAAGCTTATCGGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
               ||||||||||||||||||||||||||||||||||||||||||||||:::|:'||||:||||
b          1 AspLysLeuIleGlySerCysValTrpGlyAlaValAsnTyrThrThrAspCysAsnGly 20

y         61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
               ||||||||||||||||||||||||||:||||||||||:|||||:|||:||||||:||||||
b         21 GluCysLysArgArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnValAsnCys 40

y        121 TGGTGCGAGACT 132
             :|||||||||||
b         41 TrpCysGluThr 44

ESULT 8
BB76872
D   ABB76872 standard; Protein; 44 AA.
X
C   ABB76872;
X
T   03-JUL-2002  (first entry)
X
E   Antifungal/antibacterial peptide pEM21.
X
W   Antifungal; antibacterial; fungicide; heliomicine; human medicine;
W   veterinary medicine; plant protection; bacterial infection;
W   fungal infection.
X
S   Synthetic.
X
N   WO200206324-A2.
X
D   24-JAN-2002.
X
F   05-JUL-2001; 2001WO-FR02164.
X
R   13-JUL-2000; 2000FR-0009248.
R   19-SEP-2000; 2000FR-0011949.
X
A   (ENTO-) ENTOMED SA.
X
I   Dimarcq J,  Legrain M,  Menin L;
X
R   WPI; 2002-179779/23.
X
T   New peptide derivatives of heliomicine, useful as antibacterial and
T   antifungal agents in human or veterinary medicine and agriculture  -
X
S   Claim 15; Page 57; 100pp; French.
X
C   The present invention relates to peptide derivatives of heliomicine with
C   antifungal and/or antibacterial activity. The present sequence is one
C   such peptide. The peptides are useful in human and veterinary medicine
```

```
CC   and in plant protection, for control of bacterial and fungal infections.
CC   An antifungal and/or antibacterial composition containing at least one
CC   peptide is also claimed.
XX
SQ   Sequence   44 AA;

Alignment Scores:
Pred. No.:              1.17e-26       Length:        44
Score:                  264.00         Matches:       43
Percent Similarity:     100.00%        Conservative:  1
Best Local Similarity:  97.73%         Mismatches:    0
Query Match:            98.88%         Indels:        0
DB:                     23             Gaps:          0

US-09-673-274B-2_COPY_1_132 (1-132) x ABB76872 (1-44)

Qy          1 GATAAGCTTATCGGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
              ||'|:||||:'|||':|||':||||':||||:::|||||||||||||||||| :|||||'
Db          1 AspLysLeuIleGlySerCysValTrpGlyAlaValAsnTyrThrSerAspCysAsnGly 20

Qy         61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
              ||:'|||'||||:|·|||':||||.:||||:||||||:||||||:||||||'':||||:'' |
Db         21 GluCysLysArgArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnValAsnCys 40

Qy        121 TGGTGCGAGACT 132
              ||:|||:::'||
Db         41 TrpCysGlnThr 44

RESULT 9
ABB76896
ID   ABB76896 standard; Protein; 44 AA.
XX
AC   ABB76896;
XX
DT   03-JUL-2002  (first entry)
XX
DE   Antifungal/antibacterial peptide #3.
XX
KW   Antifungal; antibacterial; fungicide; heliomicine; human medicine;
KW   veterinary medicine; plant protection; bacterial infection;
KW   fungal infection.
XX
OS   Synthetic.
XX
PN   WO200206324-A2.
XX
PD   24-JAN-2002.
XX
PF   05-JUL-2001; 2001WO-FR02164.
XX
PR   13-JUL-2000; 2000FR-0009248.
PR   19-SEP-2000; 2000FR-0011949.
XX
PA   (ENTO-) ENTOMED SA.
XX
PI   Dimarcq J,  Legrain M,  Menin L;
XX
DR   WPI; 2002-179779/23.
XX
PT   New peptide derivatives of heliomicine, useful as antibacterial and
PT   antifungal agents in human or veterinary medicine and agriculture  -
XX
PS   Disclosure; Page 88; 100pp; French.
XX
CC   The present invention relates to peptide derivatives of heliomicine with
CC   antifungal and/or antibacterial activity. The present sequence is one
CC   such peptide. The peptides are useful in human and veterinary medicine
CC   and in plant protection, for control of bacterial and fungal infections.
CC   An antifungal and/or antibacterial composition containing at least one
CC   peptide is also claimed.
XX
SQ   Sequence   44 AA;
```

```
Alignment Scores:
Pred. No.:                 1.17e-26        Length:          44
Score:                     264.00          Matches:         43
Percent Similarity:        100.00%         Conservative:    1
Best Local Similarity:     97.73%          Mismatches:      0
Query Match:               98.88%          Indels:          0
DB:                        23              Gaps:            0

US-09-673-274B-2_COPY_1_132 (1-132) x ABB76896 (1-44)

Qy          1 GATAAGCTTATCGGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
                |||:|||||:||||||:|||||||||||:||||||||||||:|||||||:||||||||||
Db          1 AspLysLeuIleGlySerCysValTrpGlyAlaValAsnTyrThrSerAspCysAsnGly 20

Qy         61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
              |::||||||:|||||:|||||:|||||:||||:|||||:||| |||:|||||:|||:||
Db         21 GluCysLysArgArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnValAsnCys 40

Qy        121 TGGTGCGAGACT 132
              |||:|:::||:|
Db         41 TrpCysGlnThr 44

RESULT 10
ABB76862
ID   ABB76862 standard; Protein; 44 AA.
XX
AC   ABB76862;
XX
DT   03-JUL-2002  (first entry)
XX
DE   Antifungal/antibacterial peptide pEM37.
XX
KW   Antifungal; antibacterial; fungicide; heliomicine; human medicine;
KW   veterinary medicine; plant protection; bacterial infection;
KW   fungal infection.
XX
OS   Synthetic.
XX
PN   WO200206324-A2.
XX
PD   24-JAN-2002.
XX
PF   05-JUL-2001; 2001WO-FR02164.
XX
PR   13-JUL-2000; 2000FR-0009248.
PR   19-SEP-2000; 2000FR-0011949.
XX
PA   (ENTO-) ENTOMED SA.
XX
PI   Dimarcq J,  Legrain M,  Menin L;
XX
DR   WPI; 2002-179779/23.
XX
PT   New peptide derivatives of heliomicine, useful as antibacterial and
PT   antifungal agents in human or veterinary medicine and agriculture  -
XX
PS   Claim 15; Page 57; 100pp; French.
XX
CC   The present invention relates to peptide derivatives of heliomicine with
CC   antifungal and/or antibacterial activity. The present sequence is one
CC   such peptide. The peptides are useful in human and veterinary medicine
CC   and in plant protection, for control of bacterial and fungal infections.
CC   An antifungal and/or antibacterial composition containing at least one
CC   peptide is also claimed.
XX
SQ   Sequence   44 AA;

Alignment Scores:
Pred. No.:                 2.16e-26        Length:          44
Score:                     262.00          Matches:         43
Percent Similarity:        100.00%         Conservative:    1
Best Local Similarity:     97.73%          Mismatches:      0
Query Match:               98.13%          Indels:          0
DB:                        23              Gaps:            0

US-09-673-274B-2_COPY_1_132 (1-132) x ABB76862 (1-44)

Qy          1 GATAAGCTTATCGGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
              :::|||||:|||||||||||||||||||||||||||||||||||||||||||||:|:|||
Db          1 AsnLysLeuIleGlySerCysValTrpGlyAlaValAsnTyrThrSerAspCysAsnGly 20

Qy         61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
              |||||||:||||:|||||||||:||||||||||||||||||||||||||:|||||:|||
Db         21 GluCysLysArgArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnValAsnCys 40

Qy        121 TGGTGCGAGACT 132
              ||||:|||:|||
Db         41 TrpCysGluThr 44

RESULT 11
ABB76865
ID   ABB76865 standard; Protein; 44 AA.
XX
AC   ABB76865;
XX
DT   03-JUL-2002  (first entry)
XX
DE   Antifungal/antibacterial peptide pEM42.
XX
KW   Antifungal; antibacterial; fungicide; heliomicine; human medicine;
KW   veterinary medicine; plant protection; bacterial infection;
KW   fungal infection.
XX
OS   Synthetic.
XX
PN   WO200206324-A2.
XX
PD   24-JAN-2002.
XX
PF   05-JUL-2001; 2001WO-FR02164.
XX
PR   13-JUL-2000; 2000FR-0009248.
PR   19-SEP-2000; 2000FR-0011949.
XX
PA   (ENTO-) ENTOMED SA.
XX
PI   Dimarcq J,  Legrain M,  Menin L;
XX
DR   WPI; 2002-179779/23.
XX
PT   New peptide derivatives of heliomicine, useful as antibacterial and
PT   antifungal agents in human or veterinary medicine and agriculture  -
XX
PS   Claim 15; Page 57; 100pp; French.
XX
CC   The present invention relates to peptide derivatives of heliomicine with
CC   antifungal and/or antibacterial activity. The present sequence is one
CC   such peptide. The peptides are useful in human and veterinary medicine
CC   and in plant protection, for control of bacterial and fungal infections.
CC   An antifungal and/or antibacterial composition containing at least one
CC   peptide is also claimed.
XX
SQ   Sequence   44 AA;

Alignment Scores:
Pred. No.:                 2.16e-26        Length:          44
Score:                     262.00          Matches:         43
Percent Similarity:        97.73%          Conservative:    0
Best Local Similarity:     97.73%          Mismatches:      1
Query Match:               98.13%          Indels:          0
DB:                        23              Gaps:            0

US-09-673-274B-2_COPY_1_132 (1-132) x ABB76865 (1-44)

Qy          1 GATAAGCTTATCGGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
              |||::|||:|||||||||||||||||:|||||||||:|||||||||||   |||||:|||||
```

```
b          1 AspLysLeuIleGlySerCysValTrpGlyAlaValAsnTyrThrArgAspCysAsnGly 20

y         61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         21 GluCysLysArgArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnValAsnCys 40

y        121 TGGTGCGAGACT 132
             ||||||||||||
b         41 TrpCysGluThr 44

ESULT 12
BB76867
D   ABB76867 standard; Protein; 44 AA.
X
C   ABB76867;
X
T   03-JUL-2002  (first entry)
X
E   Antifungal/antibacterial peptide pEM22.
X
W   Antifungal; antibacterial; fungicide; heliomicine; human medicine;
W   veterinary medicine; plant protection; bacterial infection;
W   fungal infection.
X
S   Synthetic.
X
N   WO200206324-A2.
X
D   24-JAN-2002.
X
F   05-JUL-2001; 2001WO-FR02164.
X
R   13-JUL-2000; 2000FR-0009248.
R   19-SEP-2000; 2000FR-0011949.
X
A   (ENTO-) ENTOMED SA.
X
I   Dimarcq J,  Legrain M,  Menin L;
X
R   WPI; 2002-179779/23.
X
T   New peptide derivatives of heliomicine, useful as antibacterial and
T   antifungal agents in human or veterinary medicine and agriculture  -
X
S   Claim 15; Page 57; 100pp; French.
X
C   The present invention relates to peptide derivatives of heliomicine with
C   antifungal and/or antibacterial activity. The present sequence is one
C   such peptide. The peptides are useful in human and veterinary medicine
C   and in plant protection, for control of bacterial and fungal infections.
C   An antifungal and/or antibacterial composition containing at least one
C   peptide is also claimed.
X
Q   Sequence   44 AA;

lignment Scores:
red. No.:               2.16e-26      Length:        44
core:                   262.00        Matches:       43
ercent Similarity:      97.73%        Conservative:  0
est Local Similarity:   97.73%        Mismatches:    1
uery Match:             98.13%        Indels:        0
B:                      23            Gaps:          0

S-09-673-274B-2_COPY_1_132 (1-132) x ABB76867 (1-44)

y          1 GATAAGCTTATCGGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b          1 AspLysLeuIleGlySerCysValTrpGlyAlaValAsnTyrThrSerAspCysAsnGly 20

y         61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
             |||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
b         21 GluCysLysArgArgGlyTyrLysGlyGlyHisCysGlySerPheIleAsnValAsnCys 40
```

```
Qy       121 TGGTGCGAGACT 132
             ||||||||||||
Db        41 TrpCysGluThr 44

RESULT 13
ABB76869
ID   ABB76869 standard; Protein; 44 AA.
XX
AC   ABB76869;
XX
DT   03-JUL-2002  (first entry)
XX
DE   Antifungal/antibacterial peptide pEM25.
XX
KW   Antifungal; antibacterial; fungicide; heliomicine; human medicine;
KW   veterinary medicine; plant protection; bacterial infection;
KW   fungal infection.
XX
OS   Synthetic.
XX
PN   WO200206324-A2.
XX
PD   24-JAN-2002.
XX
PF   05-JUL-2001; 2001WO FR02164.
XX
PR   13-JUL-2000; 2000FR-0009248.
PR   19-SEP-2000; 2000FR-0011949.
XX
PA   (ENTO-) ENTOMED SA.
XX
PI   Dimarcq J,  Legrain M,  Menin L;
XX
DR   WPI; 2002-179779/23.
XX
PT   New peptide derivatives of heliomicine, useful as antibacterial and
PT   antifungal agents in human or veterinary medicine and agriculture  -
XX
PS   Claim 15; Page 57; 100pp; French.
XX
CC   The present invention relates to peptide derivatives of heliomicine with
CC   antifungal and/or antibacterial activity. The present sequence is one
CC   such peptide. The peptides are useful in human and veterinary medicine
CC   and in plant protection, for control of bacterial and fungal infections.
CC   An antifungal and/or antibacterial composition containing at least one
CC   peptide is also claimed.
XX
SQ   Sequence   44 AA;

Alignment Scores:
Pred. No.:              2.16e-26      Length:        44
Score:                  262.00        Matches:       43
Percent Similarity:     97.73%        Conservative:  0
Best Local Similarity:  97.73%        Mismatches:    1
Query Match:            98.13%        Indels:        0
DB:                     23            Gaps:          0

US-09-673-274B-2_COPY_1_132 (1-132) x ABB76869 (1-44)

Qy         1 GATAAGCTTATCGGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 AspLysLeuIleGlySerCysValTrpGlyAlaValAsnTyrThrSerAspCysAsnGly 20

Qy        61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
             |||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db        21 GluCysLysArgArgGlyTyrLysGlyGlyHisCysGlySerPheLeuAsnValAsnCys 40

Qy       121 TGGTGCGAGACT 132
             ||||||||||||
Db        41 TrpCysGluThr 44

RESULT 14
ABB76871
```

```
ID   ABB76871 standard; Protein; 44 AA.
XX
AC   ABB76871;
XX
DT   03-JUL-2002  (first entry)
XX
DE   Antifungal/antibacterial peptide pEM7.
XX
KW   Antifungal; antibacterial; fungicide; heliomicine; human medicine;
KW   veterinary medicine; plant protection; bacterial infection;
KW   fungal infection.
XX
OS   Synthetic.
XX
PN   WO200206324 A2.
XX
PD   24-JAN-2002.
XX
PF   05-JUL-2001; 2001WO-FR02164.
XX
PR   13-JUL-2000; 2000FR-0009248.
PR   19-SEP-2000; 2000FR-0011949.
XX
PA   (ENTO ) ENTOMED SA.
XX
PI   Dimarcq J,  Legrain M,  Menin L;
XX
DR   WPI; 2002-179779/23.
XX
PT   New peptide derivatives of heliomicine, useful as antibacterial and
PT   antifungal agents in human or veterinary medicine and agriculture  -
XX
PS   Claim 19; Page 57; 100pp; French.
XX
CC   The present invention relates to peptide derivatives of heliomicine with
CC   antifungal and/or antibacterial activity. The present sequence is one
CC   such peptide. The peptides are useful in human and veterinary medicine
CC   and in plant protection, for control of bacterial and fungal infections.
CC   An antifungal and/or antibacterial composition containing at least one
CC   peptide is also claimed.
XX
SQ   Sequence   44 AA;

Alignment Scores:
Pred. No.:                2.16e-26      Length:        44
Score:                    262.00        Matches:       43
Percent Similarity:       100.00%       Conservative:  0
Best Local Similarity:    100.00%       Mismatches:    0
Query Match:              98.13%        Indels:        0
DB:                       23            Gaps:          0

US-09-673-274B-2_COPY_1_132 (1-132) x ABB76871 (1-44)

Qy          1 GATAAGCTTATCGGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AspLysLeuIleGlySerCysValTrpGlyAlaValAsnTyrThrSerAspCysAsnGly 20

Qy         61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 GluCysLysArgArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnValAsnCys 40

Qy        121 TGGTGCGAG 129
              |||||||||
Db         41 TrpCysGlu 43

RESULT 15
AAY54376
ID   AAY54376 standard; Peptide; 74 AA.
XX
AC   AAY54376;
XX
DT   06-APR-2000  (first entry)
XX
DE   A fusion peptide of PR-1alpha signal peptide/heliomycin.
XX
KW   PR-1alpha; heliomycin; polycysteine peptide; antifungal; antibacterial;
KW   fungal infection; plant protection; animal protection; transgenic plant;
KW   fungi; bacteria; Cercospora beticola; Cladosporium herbarum;
KW   Fusarium culmorum; F. oxysporum; Phytophthora cinnamoni.
XX
OS   Chimeric - Nicotiana sp.
OS   Chimeric - Heliothis virescens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..30
FT                   /note= "tobacco PR-1alpha signal peptide"
FT   Peptide         31..74
FT                   /note= "heliomycin"
XX
PN   FR2777568-A1.
XX
PD   22-OCT-1999.
XX
PF   15-APR-1998;   98FR-0004933.
XX
PR   15-APR-1998;   98FR-0004933.
XX
PA   (RHON ) RHONE-POULENC AGROCHIMIE.
XX
DR   WPI; 2000-108532/10.
DR   N-PSDB; AAZ45693.
XX
PT   New polycysteine peptides, designated heliomycin, with antifungal or
PT   antibacterial activity, useful in medicine or plant protection  -
XX
PS   Claim 13; Page 32; 46pp; French.
XX
CC   The present sequence represents a fusion peptide of the tobacco
CC   PR-1alpha signal peptide and heliomycin. The attachment of the
CC   PR-1alpha sequence to the N-terminal aids secretion of the heliomycin
CC   peptide. Heliomycin is a polycysteine peptide with antifungal and
CC   antibacterial activity, isolated from the lepidopteron Heliothis
CC   virescens. Heliomycin peptides contain the sequence given in
CC   AAY54378. The heliomycin peptides are used as pharmaceuticals for
CC   treating or preventing fungal infections in humans and animals. They
CC   may also be used as antifungal agents for plant protection. Nucleic
CC   acid encoding the heliomycin peptides is used to generate transgenic
CC   plants that are resistant to some fungi and bacteria (specifically
CC   Cercospora beticola, Cladosporium herbarum, Fusarium culmorum,
CC   F. oxysporum and Phytophthora cinnamoni), and for recombinant production
CC   of the peptides.
XX
SQ   Sequence   74 AA;

Alignment Scores:
Pred. No.:                2.32e-26      Length:        74
Score:                    262.00        Matches:       44
Percent Similarity:       100.00%       Conservative:  0
Best Local Similarity:    100.00%       Mismatches:    0
Query Match:              98.13%        Indels:        0
DB:                       21            Gaps:          0

US-09-673-274B-2_COPY_1_132 (1-132) x AAY54376 (1-74)

Qy          1 GATAAGCTTATCGGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         31 AspLysLeuIleGlySerCysValTrpGlyAlaValAsnTyrThrSerAspCysAsnGly 50

Qy         61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         51 GluCysLysArgArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnValAsnCys 70

Qy        121 TGGTGCGAGACT 132
              ||||||||||||
Db         71 TrpCysGluThr 74
```

```
earch completed: October 16, 2003, 17:10:36
ob time : 41.7841 secs
```

```
                     GenCore version 5.1.6
             Copyright (c) 1993 - 2003  Compugen Ltd.

M nucleic - protein search, using frame_plus_n2p model

un on:          October 16, 2003, 17:06:33 ; Search time 11.2292 Seconds
                                             (without alignments)
                                             1415.504 Million cell updates/sec

itle:           US-09-673-274B-2
erfect score:   324
equence:        1 gataagcttatcggttcctg..........gaacgtgtcgacggatccgg 169

coring table:   BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

earched:        127863 seqs, 47026705 residues

otal number of hits satisfying chosen parameters:      255726

inimum DB seq length: 0
aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%
                Maximum Match 100%
                Listing first 45 summaries

ommand line parameters:
MODEL=frame+_n2p.model -DEV=xlp
Q=/cgn2_1/USPTO_spool/US09673274/runat_16102003_170627_2531/app_query.fasta_1.654
DB=SwissProt_41 -QFMT=fastan -SUFFIX=rsp -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
USER=US09673274_@CGN_1_1_36_@runat_16102003_170627_2531 -NCPU=6 -ICPU=3
NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

atabase :       SwissProt_41:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
                  %
esult             Query
 No.    Score   Match Length DB ID                    Description
-------------------------------------------------------------------------------
   1      267    82.4     44  1 DEFN_HELVI             P81544 heliothis v
   2       86    26.5     66  1 SCX3_BUTOM             P13488 buthus occi
   3     84.5    26.1     85  1 SCN2_MESMA             Q9bkj1 mesobuthus
   4       84    25.9     85  1 SCAS_MESMA             Q9uac9 mesobuthus
   5     83.5    25.8     85  1 SCN3_MESMA             Q9bkj0 mesobuthus
   6     79.5    24.5     66  1 SCXC_CENLL             P45667 centruroide
   7       79    24.4     85  1 SCXE_BUTOC             Q17254 buthus occi
   8     78.5    24.2     65  1 SCX4_BUTOM             P59354 buthus occi
   9     78.5    24.2     84  1 SCX3_ANDAU             P01480 androctonus
  10     77.5    23.9     64  1 SCL6_LEIQH             P59356 leiurus qui
  11     77.5    23.9     65  1 SCX4_LEIQU             P01489 leiurus qui
  12       77    23.8     85  1 SCAT_MESMA             Q9uac8 mesobuthus
  13     76.5    23.6     61  1 SIX2_LEIQU             P19855 leiurus qui
  14       76    23.5     66  1 SCXE_BUTEU             P09982 buthus eupe
  15       76    23.5     67  1 SCL3_LEIQH             P56678 leiurus qui
  16     75.5    23.3     84  1 SCXA_MESMA             O61705 mesobuthus
  17       75    23.1     84  1 SCX4_ANDAU             P45658 androctonus
  18     74.5    23.0     70  1 DMYC_DROME             P41964 drosophila
  19     74.5    23.0     87  1 SCX1_CENSC             P01492 centruroide
  20       74    22.8     59  1 SCX5_CENSC             P58779 centruroide
  21       73    22.5     61  1 SIX5_ANDAU             P81504 androctonus
  22       73    22.5     64  1 SCX5_LEIQU             P01481 leiurus qui
  23       72    22.2     85  1 SC15_MESMA             Q9gng8 mesobuthus
  24       72    22.2     85  1 SCA1_MESMA             Q9gyx2 mesobuthus
  25     71.5    22.1     61  1 SIX2_BUTAR             P80962 buthacus ar
  26     71.5    22.1     61  1 SIX4_BUTOC             P55903 buthus occi
  27     71.5    22.1     61  1 SIX5_BUTOC             P55904 buthus occi
  28     71.5    22.1     87  1 SCX2_CENSC             P01493 centruroide
  29     71.5    22.1     87  1 SCX3_CENSC             P01494 centruroide
  30       71    21.9     43  1 DEFI_PALPR             P80407 palomena pr
  31       71    21.9     66  1 SCL7_LEIQH             P59357 leiurus qui
  32     70.5    21.8     64  1 SCM2_MESMA             P58488 mesobuthus
  33     70.5    21.8     84  1 SC16_MESMA             Q9gqv6 mesobuthus
  34     70.5    21.8     87  1 SCX6_CENNO             P45664 centruroide
  35     69.5    21.5     84  1 SCX1_MESMA             P45697 mesobuthus
  36     69.5    21.5     84  1 SCX2_CENNO             P01495 centruroide
  37     69.5    21.5     84  1 SCX9_CENSC             Q95wc9 centruroide
  38     69.5    21.5     85  1 SC10_MESMA             Q9njc5 mesobuthus
  39     68.5    21.1     62  1 SIX4_BUTSI             P82814 buthus sind
  40     68.5    21.1     65  1 SCX1_BUTOC             P01488 buthus occi
  41     68.5    21.1     66  1 SCX6_ANDAU             P56743 androctonus
  42     68.5    21.1     66  1 SCX9_BUTEU             P09981 buthus eupe
  43     68.5    21.1     74  1 DEF2_CAPAN             O65740 capsicum an
  44     68.5    21.1     74  1 SC17_MESMA             Q9njc4 mesobuthus
  45     68.5    21.1     83  1 SCX1_ANDAU             P01479 androctonus


                              ALIGNMENTS

RESULT 1
DEFN_HELVI
ID   DEFN_HELVI     STANDARD;      PRT;    44 AA.
AC   P81544;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Defensin heliomicin.
OS   Heliothis virescens (Noctuid moth) (Owlet moth).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea;
OC   Noctuidae; Heliothinae; Heliothis.
OX   NCBI_TaxID=7102;
RN   [1]
RP   SEQUENCE, AND CHARACTERIZATION.
RC   TISSUE=Hemolymph;
RX   MEDLINE=99194775; PubMed=10092609;
RA   Lamberty M., Ades S., Uttenweiler-Joseph S., Brookhart G., Bushey D.,
RA   Hoffmann J.A., Bulet P.;
RT   "Insect immunity. Isolation from the lepidopteran Heliothis virescens
RT   of a novel insect defensin with potent antifungal activity.";
RL   J. Biol. Chem. 274:9320-9326(1999).
CC   -!- FUNCTION: THIS PEPTIDE HAS POTENT ANTI-FUNGAL ACTIVITY. HAS NO
CC       ACTIVITY AGAINST GRAM-NEGATIVE AND GRAM-POSITIVE BACTERIA.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- MASS SPECTROMETRY: MW=4784.8; METHOD=MALDI.
DR   PDB; 1I2U; 12-FEB-02.
DR   PDB; 1I2V; 12-FEB-02.
KW   Insect immunity; Fungicide; 3D-structure.
FT   DISULFID      7     32
FT   DISULFID     18     40
FT   DISULFID     22     42
SQ   SEQUENCE   44 AA;  4790 MW;  3C1DA1BD3376FCD5 CRC64;

Alignment Scores:
Pred. No.:              3.17e-25      Length:          44
Score:                  267.00        Matches:         44
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            82.41%        Indels:          0
DB:                     1             Gaps:            0
```

```
US-09-673-274B-2 (1-169) x DEFN_HELVI (1-44)

Qy          1 GATAAGCTTATCGGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
              :| |||||||||||||||||||:||||.|||||:.|||.:|||.:|||||||||||||||||
Db          1 AspLysLeuIleGlySerCysValTrpGlyAlaValAsnTyrThrSerAspCysAsnGly 20

Qy         61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
              |':||||||:||||::||||':|||||. |||::|||||:||||:|||:'|||||||::||
Db         21 GluCysLysArgArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnValAsnCys 40

Qy        121 TGGTGCGAGACT 132
              :::||||:||||
Db         41 TrpCysGluThr 44

RESULT 2
SCX3_BUTOM
ID   SCX3_BUTOM     STANDARD;      PRT;    66 AA.
AC   P13488;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Alpha-like neurotoxin III (Bom III) (Bom3).
OS   Buthus occitanus mardochei (Moroccan scorpion).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Buthus.
OX   NCBI_TaxID=6869;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Venom;
RX   MEDLINE=87161829; PubMed=3104036;
RA   Vargas O., Martin M.-F., Rochat H.;
RT   "Characterization of six toxins from the venom of the Moroccan
RT   scorpion Buthus occitanus mardochei.";
RL   Eur. J. Biochem. 162:589-599(1987).
CC   -!- FUNCTION: Binds to sodium channels and inhibits the inactivation
CC       of the activated channels, thereby blocking neuronal transmission.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC       ALPHA-TOXIN SUBFAMILY. ALPHA-LIKE TOXIN SUBCLASS.
DR   PIR; A29386; A29386.
DR   HSSP; P56678; 1BMR.
DR   InterPro; IPR003614; Knot1.
DR   InterPro; IPR002061; Scorpion_toxinL.
DR   Pfam; PF00537; toxin_3; 1.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor.
FT   DISULFID     12     65       BY SIMILARITY.
FT   DISULFID     16     37       BY SIMILARITY.
FT   DISULFID     23     47       BY SIMILARITY.
FT   DISULFID     27     49       BY SIMILARITY.
SQ   SEQUENCE   66 AA;  6872 MW;  7A96DA393F89852F CRC64;

Alignment Scores:
Pred. No.:                0.00893       Length:        66
Score:                    86.00         Matches:       15
Percent Similarity:       55.00%        Conservative:  7
Best Local Similarity:    37.50%        Mismatches:    16
Query Match:              26.54%        Indels:        2
DB:                       1             Gaps:          1

US-09-673-274B-2 (1-169) x SCX3_BUTOM (1-66)

Qy         16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
              :::|||':||:::                     :::::||   ||||:::         ||||:.|    :::
Db         11 AsnCysValTyrHisCysPheProGlySerSerGlyCysAspThrLeuCysLysGluLys 30

Qy         76 GGTTACAAGGGTGGTCACTGCGGTTCCTTC------GCTAACGTGAACTGCTGGTGCGAG 129
               :||          |:||||||.||||           :::    |||    ||||'||||::::
Db         31 GlyAlaThrSerGlyHisCysGlyPheLeuProGlySerGlyValAlaCysTrpCysAsp 50

RESULT 3
```

```
SCN2_MESMA
ID   SCN2_MESMA     STANDARD;      PRT;    85 AA.
AC   Q9BKJ1;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Anti-neuroexcitation peptide II precursor (ANEPII).
OS   Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Mesobuthus.
OX   NCBI_TaxID=34649;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Venom gland;
RA   Zhang J.-H., Hua Z.C., Zhu D.X.;
RT   "Cloning of anti-neuroexcitation peptide II (ANEP) cDNA from Scorpion
RT   Buthus martensii Karsch.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Binds to sodium channels and inhibits them.
CC   -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC   -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC       ALPHA-TOXIN SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF242736; AAK28341.1; -.
DR   HSSP; P01491; 1B3C.
DR   InterPro; IPR003614; Knot1.
DR   InterPro; IPR002061; Scorpion_toxinL.
DR   Pfam; PF00537; toxin_3; 1.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor;
KW   Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22     85       ANTI-NEUROEXCITATION PEPTIDE II.
FT   DISULFID     31     81       BY SIMILARITY.
FT   DISULFID     35     56       BY SIMILARITY.
FT   DISULFID     42     63       BY SIMILARITY.
FT   DISULFID     46     65       BY SIMILARITY.
SQ   SEQUENCE   85 AA;  9224 MW;  C045C7DACB3513AB CRC64;

Alignment Scores:
Pred. No.:                0.014         Length:        85
Score:                    84.50         Matches:       15
Percent Similarity:       55.26%        Conservative:  6
Best Local Similarity:    39.47%        Mismatches:    12
Query Match:              26.08%        Indels:        5
DB:                       1             Gaps:          2

US-09-673-274B-2 (1-169) x SCN2_MESMA (1-85)

Qy         16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
              ||||||:::||||||                                   ||||||    ||||||:::
Db         34 SerCysLeuTrpGly------------AsnAspGlyCysAsnLysGluCysArgAlaTyr 49

Qy         76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
              |||          ||:::|||   ::::::        :::   ||||||||||||
Db         50 GlyAlaSerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 66

RESULT 4
SCAS_MESMA
ID   SCAS_MESMA     STANDARD;      PRT;    85 AA.
AC   Q9UAC9;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
```

```
T     28-FEB-2003 (Rel. 41, Last annotation update)
E     Neurotoxin AS precursor.
S     Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).
C     Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
C     Buthoidea; Buthidae; Mesobuthus.
X     NCBI_TaxID=34649;
N     [1]
P     SEQUENCE FROM N.A.
C     TISSUE=Venom gland;
A     Lan Z.D., Dai L., Chi C.-W.;
T     "The gene cloning and sequencing of AS and AS-1, two novel neurotoxins
T     from the scorpion Buthus martensi Karsch.";
L     Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
C     -!- FUNCTION: Ryanodine receptor agonist in skeletal muscle.
C     -!- SUBCELLULAR LOCATION: Secreted (By similarity).
C     -!- TISSUE SPECIFICITY: Expressed by the venom gland.
C     -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
C         ALPHA-TOXIN SUBFAMILY.
C     --------------------------------------------------------------------------
C     This SWISS-PROT entry is copyright. It is produced through a collaboration
C     between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
C     the European Bioinformatics Institute.  There are no  restrictions on  its
C     use  by  non-profit  institutions as long  as its content  is  in  no  way
C     modified and this statement is not removed.  Usage  by  and for commercial
C     entities requires a license agreement (See http://www.isb-sib.ch/announce/
C     or send an email to license@isb-sib.ch).
C     --------------------------------------------------------------------------
R     EMBL; AF079060; AAD47374.1; -.
R     HSSP; P01492; 1VNB.
R     InterPro; IPR003614; Knot1.
R     InterPro; IPR002061; Scorpion_toxinL.
R     Pfam; PF00537; toxin_3; 1.
R     ProDom; PD000908; Scorpion_toxinL; 1.
R     SMART; SM00505; Knot1; 1.
W     Toxin; Neurotoxin; Signal.
T     SIGNAL        1     19       POTENTIAL.
T     CHAIN        20     85       NEUROTOXIN AS.
T     DISULFID     31     81       BY SIMILARITY.
T     DISULFID     35     56       BY SIMILARITY.
T     DISULFID     42     63       BY SIMILARITY.
T     DISULFID     46     65       BY SIMILARITY.
Q     SEQUENCE   85 AA;  9759 MW;  41286F9DDCD4FE90 CRC64;

lignment Scores:
red. No.:                0.0161         Length:         85
core:                    84.00          Matches:        19
ercent Similarity:       55.81%         Conservative:   5
est Local Similarity:    44.19%         Mismatches:     17
uery Match:              25.93%         Indels:         2
B:                       1              Gaps:           2

S-09-673-274B-2 (1-169) x SCAS_MESMA (1-85)

y           1 GATAAGCTTATCGGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
              ||||||      |||      ||||||      :::|||      |:|   ||||||
b          26 AspLysTyrThrGlyCysLysValTrpCysValIleAsnAsnGluSer---CysAsnSer 44

y          61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
              |||||||||   ||||||         |||:::|||      |||      :::   |||
b          45 GluCysLysIleArgGlyGlyTyrTyrGlyTyrCys---TyrPheTrpLysLeuAlaCys 63

y         121 TGGTGCGAG 129
              :::|||:::
b          64 PheCysGln 66

RESULT 5
CN3_MESMA
D     SCN3_MESMA     STANDARD;      PRT;    85 AA.
C     Q9BKJ0; Q9GZC4;
T     28-FEB-2003 (Rel. 41, Created)
T     28-FEB-2003 (Rel. 41, Last sequence update)
T     28-FEB-2003 (Rel. 41, Last annotation update)
E     Anti-neuroexcitation peptide III precursor (ANEPIII) (Anti-epilepsy
```

```
DE    peptide).
OS    Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).
OC    Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC    Buthoidea; Buthidae; Mesobuthus.
OX    NCBI_TaxID=34649;
RN    [1]
RP    SEQUENCE FROM N.A.
RC    TISSUE=Venom gland;
RA    Zhang J.-H., Hua Z.C., Zhu D.X.;
RT    "Cloning of anti-epilepsy peptide cDNA from scorpion Buthus martensii
RT    Karsch.";
RL    Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
RN    [2]
RP    SEQUENCE FROM N.A.
RC    TISSUE=Venom gland;
RA    Zhang J.-H., Hua Z.C., Zhu D.X.;
RT    "Cloning of anti-neuroexcitation peptide III (ANEP) cDNA from Scorpion
RT    Buthus martensii Karsch.";
RL    Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
CC    -!- FUNCTION: Binds to sodium channels and inhibits them.
CC    -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC    -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC    -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC        ALPHA-TOXIN SUBFAMILY.
CC    --------------------------------------------------------------------------
CC    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC    between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC    the European Bioinformatics Institute.  There are no  restrictions on  its
CC    use  by  non-profit  institutions as long  as its content  is  in  no  way
CC    modified and this statement is not removed.  Usage  by  and for commercial
CC    entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC    or send an email to license@isb-sib.ch).
CC    --------------------------------------------------------------------------
DR    EMBL; AF122003; AAG01571.1; -.
DR    EMBL; AF242737; AAK28342.1; -.
DR    HSSP; P01494; 2SN3.
DR    InterPro; IPR003614; Knot1.
DR    InterPro; IPR002061; Scorpion_toxinL.
DR    Pfam; PF00537; toxin_3; 1.
DR    ProDom; PD000908; Scorpion_toxinL; 1.
DR    SMART; SM00505; Knot1; 1.
KW    Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor;
KW    Signal.
FT    SIGNAL        1     21       POTENTIAL.
FT    CHAIN        22     85       ANTI-NEUROEXCITATION PEPTIDE III.
FT    DISULFID     31     81       BY SIMILARITY.
FT    DISULFID     35     56       BY SIMILARITY.
FT    DISULFID     42     63       BY SIMILARITY.
FT    DISULFID     46     65       BY SIMILARITY.
FT    CONFLICT     47     47       K -> I (IN REF. 2).
SQ    SEQUENCE   85 AA;  9271 MW;  C1D6C93A2F82F8C2 CRC64;

Alignment Scores:
Pred. No.:                0.0186         Length:         85
Score:                    83.50          Matches:        16
Percent Similarity:       55.26%         Conservative:   5
Best Local Similarity:    42.11%         Mismatches:     12
Query Match:              25.77%         Indels:         5
DB:                       1              Gaps:           2

US-09-673-274B-2 (1-169) x SCN3_MESMA (1-85)

Qy         16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
              ||||||:::||||||                     ||||||   ||||||:||
Db         34 SerCysLeuTrpGly------------AsnGluGlyCysAsnLysGluCysLysGlyPhe 49

Qy         76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
              |||         |||:::|||   ::::::      :::   ||||||||.|:|
Db         50 GlyAlaTyrTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 66

RESULT 6
SCXC_CENLL
ID    SCXC_CENLL     STANDARD;      PRT;    66 AA.
```

```
AC   P45667;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Crustacean-specific toxin 1.
OS   Centruroides limpidus limpidus (Mexican scorpion).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Centruroides.
OX   NCBI_TaxID=29941;
RN   [1]
RP   SEQUENCE, DISULFIDE BONDS, AND STRUCTURE BY NMR.
RC   TISSUE=Venom;
RX   MEDLINE=95244424; PubMed=7727365;
RA   Lebreton F., Delepierre M., Ramirez A.N., Balderas C., Possani L.D.;
RT   "Primary and NMR three-dimensional structure determination of a novel
RT   crustacean toxin from the venom of the scorpion Centruroides limpidus
RT   limpidus Karsch.";
RL   Biochemistry 33:11135-11149(1994).
CC   -!- FUNCTION: Binds to sodium channels and inhibits the inactivation
CC       of the activated channels, thereby blocking neuronal transmission.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC       BETA-TOXIN SUBFAMILY.
DR   PIR; A55869; A55869.
DR   HSSP; P01492; 1VNB.
DR   InterPro; IPR003614; Knot1.
DR   InterPro; IPR001219; Neurotoxin.
DR   InterPro; IPR002061; Scorpion_toxinL.
DR   Pfam; PF00537; toxin_3; 1.
DR   PRINTS; PR00284; TOXIN.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor.
FT   DISULFID     12     65
FT   DISULFID     16     41
FT   DISULFID     25     46
FT   DISULFID     29     48
SQ   SEQUENCE   66 AA;  7338 MW;  A478E63796F54DAA CRC64;

Alignment Scores:
Pred. No.:                0.0571        Length:        66
Score:                    79.50         Matches:       17
Percent Similarity:       53.85%        Conservative:  4
Best Local Similarity:    43.59%        Mismatches:    11
Query Match:              24.54%        Indels:        7
DB:                       1             Gaps:          3

US-09-673-274B-2 (1-169) x SCXC_CENLL (1-66)

Qy          19 TGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGGGGT 78
               |||   |||         |||         :::|||:::   |||||||||   :::
Db          16 CysPheTrpLeuGlyLysAsn------GluAsnCysAspLysGluCysLysAlaLysAsn 33

Qy          79 TACAAGGGT------GGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
                     |||      |||:::|||   ||||.||||         ||||||||||||
Db          34 GlnGlyGlySerTyrGlyTyrCysTyrSerPheAla---------CysTrpCysGlu 49

RESULT 7
SCXE_BUTOC
ID   SCXE_BUTOC      STANDARD;      PRT;    85 AA.
AC   Q17254;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Neurotoxin XIV precursor (Bot14).
OS   Buthus occitanus tunetanus (Common European scorpion).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Buthus.
OX   NCBI_TaxID=6871;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Venom;
RX   MEDLINE=96300228; PubMed=8706664;
RA   Bouhaouala-Zahar B., Ducancel F., Zenouaki I., Ben Khalifa R.,
RA   Borchani L., Pelhate M., Boulain J.C., el Ayeb M., Menez A.,
RA   Karoui H.;
RT   "A recombinant insect-specific alpha-toxin of Buthus occitanus
RT   tunetanus scorpion confers protection against homologous mammal
RT   toxins.";
RL   Eur. J. Biochem. 238:653-660(1996).
CC   -!- FUNCTION: Binds to sodium channels and inhibits the inactivation
CC       of the activated channels, thereby blocking neuronal transmission.
CC       This toxin is active only on insects.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC       ALPHA-TOXIN SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X92376; CAA63120.1; -.
DR   PIR; S68906; S68906.
DR   HSSP; P56678; 1BMR.
DR   InterPro; IPR003614; Knot1.
DR   InterPro; IPR002061; Scorpion_toxinL.
DR   Pfam; PF00537; toxin_3; 1.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor;
KW   Signal.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19     85       NEUROTOXIN XIV.
FT   DISULFID     30     83       BY SIMILARITY.
FT   DISULFID     34     55       BY SIMILARITY.
FT   DISULFID     41     65       BY SIMILARITY.
FT   DISULFID     45     67       BY SIMILARITY.
SQ   SEQUENCE   85 AA;  9174 MW;  4977F16C511F0567 CRC64;

Alignment Scores:
Pred. No.:                0.0671        Length:        85
Score:                    79.00         Matches:       14
Percent Similarity:       50.00%        Conservative:  6
Best Local Similarity:    35.00%        Mismatches:    18
Query Match:              24.38%        Indels:        2
DB:                       1             Gaps:          1

US-09-673-274B-2 (1-169) x SCXE_BUTOC (1-85)

Qy          16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
               :::|||   :::      :::      :::|||   |||:::      ||||||
Db          29 AsnCysAlaTyrHisCysLeuLysIleSerSerGlyCysAspThrLeuCysLysGluAsn 48

Qy          76 GGTTACAAGGGTGGTCACTGCGGT------TCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
               |||         ||||||||||||      |||               |||||||||:::
Db          49 GlyAlaThrSerGlyHisCysGlyHisLysSerGlyHisGlySerAlaCysTrpCysLys 68

RESULT 8
SCX4_BUTOM
ID   SCX4_BUTOM      STANDARD;      PRT;    65 AA.
AC   P59354;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Alpha-like neurotoxin IV (Bom IV) (Bom4).
OS   Buthus occitanus mardochei (Moroccan scorpion).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Buthus.
OX   NCBI_TaxID=6869;
RN   [1]
```

```
P   SEQUENCE, PHARMACOLOGICAL CHARACTERIZATION, AND MASS SPECTROMETRY.
C   TISSUE=Venom;
X   MEDLINE=99203324; PubMed=10103091;
A   Cestele S., Stankiewicz M., Mansuelle P., De Waard M., Dargent B.,
A   Gilles N., Pelhate M., Rochat H., Martin-Eauclaire M.-F., Gordon D.;
T   "Scorpion alpha-like toxins, toxic to both mammals and insects,
T   differentially interact with receptor site 3 on voltage-gated sodium
T   channels in mammals and insects.";
L   Eur. J. Neurosci. 11:975-985(1999).
C   -!- FUNCTION: Binds to sodium channels and inhibits the inactivation
C       of the activated channels, thereby blocking neuronal transmission.
C   -!- SUBCELLULAR LOCATION: Secreted.
C   -!- TISSUE SPECIFICITY: Expressed by the venom gland.
C   -!- MASS SPECTROMETRY: MW=7287.96; MW_ERR=0.37; METHOD=Electrospray.
C   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
C       ALPHA-TOXIN SUBFAMILY. ALPHA-LIKE TOXIN SUBCLASS.
R   Pfam; PF00537; toxin_3; 1.
R   ProDom; PD000908; Scorpion_toxinL; 1.
R   SMART; SM00505; Knot1; 1.
W   Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor.
T   DISULFID     12     63        BY SIMILARITY.
T   DISULFID     16     36        BY SIMILARITY.
T   DISULFID     22     46        BY SIMILARITY.
T   DISULFID     26     48        BY SIMILARITY.
Q   SEQUENCE   65 AA;  7296 MW;  4305EB32436A4B90 CRC64;

lignment Scores:
red. No.:              0.0759         Length:        65
core:                  78.50          Matches:       16
ercent Similarity:     51.22%         Conservative:  5
est Local Similarity:  39.02%         Mismatches:    15
uery Match:            24.23%         Indels:        5
B:                     1              Gaps:          3

S-09-673-274B-2 (1-169) x SCX4_BUTOM (1-65)

y         16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
             :::||||||:::                  |||   ||||||      ||:   :::
b         11 AsnCysValTyrGluCysAlaLys---AsnSerTyrCysAsnAspLeuCysThrLysAsn 29

y         76 GGTTACAAGGGTGGTCACTGC---------GGTTCCTTCGCTAACGTGAACTGCTGGTGC 126
             |||   |||   |||:::|||         |||   :::   ||.      |||||||||
b         30 GlyAlaLysSerGlyTyrCysGlnTrpLeuGlyLysTyrGlyAsnAla---CysTrpCys 48

y        127 GAG 129
             |||
b         49 Glu 49

ESULT 9
CX3_ANDAU
D   SCX3_ANDAU     STANDARD;      PRT;    84 AA.
C   P01480;
T   21-JUL-1986 (Rel. 01, Created)
T   01-JUL-1993 (Rel. 26, Last sequence update)
T   28-FEB-2003 (Rel. 41, Last annotation update)
E   Neurotoxin III precursor (AaH III) (AaH3).
S   Androctonus australis hector (Sahara scorpion).
C   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
C   Buthoidea; Buthidae; Androctonus.
X   NCBI_TaxID=70175;
N   [1]
P   SEQUENCE FROM N.A.
X   MEDLINE=90037062; PubMed=2808423;
A   Bougis P.E., Rochat H., Smith L.A.;
T   "Precursors of Androctonus australis scorpion neurotoxins. Structures
T   of precursors, processing outcomes, and expression of a functional
T   recombinant toxin II.";
L   J. Biol. Chem. 264:19259-19265(1989).
N   [2]
P   SEQUENCE OF 20-83.
X   MEDLINE=79148414; PubMed=428402;
A   Kopeyan C., Martinez G., Rochat H.;
T   "Amino acid sequence of neurotoxin III of the scorpion Androctonus
```

```
RT   australis Hector.";
RL   Eur. J. Biochem. 94:609-615(1979).
RN   [3]
RP   STRUCTURE BY NMR.
RX   MEDLINE=91122921; PubMed=2279844;
RA   Laplante S.R., Mikou A., Robin M., Guittet E., Delsuc M.A.,
RA   Charpentier I., Lallemand J.-Y.;
RT   "Rapid determination and NMR assignments of antiparallel sheets and
RT   helices of a scorpion and a cobra toxin.";
RL   Int. J. Pept. Protein Res. 36:227-230(1990).
RN   [4]
RP   STRUCTURE BY NMR.
RX   MEDLINE=93044377; PubMed=1422146;
RA   Mikou A., Laplante S.R., Guittet E., Lallemand J.-Y.,
RA   Martin-Eauclaire M.-F., Rochat H.;
RT   "Toxin III of the scorpion Androctonus australis hector: proton
RT   nuclear magnetic resonance assignments and secondary structure.";
RL   J. Biomol. NMR 2:57-70(1992).
CC   -!- FUNCTION: Binds to sodium channels and inhibits the inactivation
CC       of the activated channels, thereby blocking neuronal transmission.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC       ALPHA-TOXIN SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M27703; AAA29948.1; -.
DR   PIR; C34444; NTSR3A.
DR   HSSP; P56678; 1BMR.
DR   InterPro; IPR003614; Knot1.
DR   InterPro; IPR002061; Scorpion_toxinL.
DR   Pfam; PF00537; toxin_3; 1.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor;
KW   Signal.
FT   SIGNAL        1     19
FT   CHAIN        20     83       NEUROTOXIN III.
FT   PROPEP       84     84
FT   DISULFID     31     81       BY SIMILARITY.
FT   DISULFID     35     53       BY SIMILARITY.
FT   DISULFID     39     63       BY SIMILARITY.
FT   DISULFID     43     65       BY SIMILARITY.
FT   CONFLICT     27     27       D -> N (IN REF. 2).
SQ   SEQUENCE   84 AA;  9068 MW;  ADB0AD46215 3E3CC CRC64;

Alignment Scores:
Pred. No.:              0.0773         Length:        84
Score:                  78.50          Matches:       15
Percent Similarity:     53.85%         Conservative:  6
Best Local Similarity:  38.46%         Mismatches:    13
Query Match:            24.23%         Indels:        5
DB:                     1              Gaps:          2

US-09-673-274B-2 (1-169) x SCX3_ANDAU (1-84)

Qy        16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
             :::'|||.':::      |||               |||:::||.   || |||:::
Db        30 AsnCysValTyrHisCysVal---------ProProCysAspGlyLeuCysLysLysAsn 46

Qy        76 GGTTACAAGGGTGGTCACTGCGGTTCCTTC------GCTAACGTGAACTGCTGGTGC 126
             |||   ||'   |||   ''||||            :::   :::   ||.|||| |
Db        47 GlyAlaLysSerGlySerCysGlyPheLeuIleProSerGlyLeuAlaCysTrpCys 65

RESULT 10
SCL6_LEIQH
```

```
ID   SCL6_LEIQH     STANDARD;      PRT;    64 AA.
AC   P59356;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Alpha-like neurotoxin Lqh VI (Lqh6).
OS   Leiurus quinquestriatus hebraeus (Yellow scorpion).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Leiurus.
OX   NCBI_TaxID=6884;
RN   [1]
RP   SEQUENCE, PHARMACOLOGICAL CHARACTERIZATION, AND BIOASSAY.
RC   TISSUE=Venom;
RX   MEDLINE=22168983; PubMed=12180969;
RA   Hamon A., Gilles N., Sautiere P., Martinage A., Kopeyan C., Ulens C.,
RA   Tytgat J., Lancelin J.-M., Gordon D.;
RT   "Characterization of scorpion alpha-like toxin group using two new
RT   toxins from the scorpion Leiurus quinquestriatus hebraeus.";
RL   Eur. J. Biochem. 269:3920-3933(2002).
CC   -!- FUNCTION: Binds to sodium channels and inhibits the inactivation
CC       of the activated channels, thereby blocking neuronal transmission.
CC       This toxin is highly toxic to insects and mice, and inhibits the
CC       binding of alpha-toxin to cockroach neuronal membranes.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC   -!- MISCELLANEOUS: LD(50) is 34.3 nmol/kg to cockroaches (Blatella
CC       germanica).
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC       ALPHA-TOXIN SUBFAMILY. ALPHA-LIKE TOXIN SUBCLASS.
DR   Pfam; PF00537; toxin_3; 1.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor;
KW   Amidation.
FT   DISULFID     12     62        BY SIMILARITY.
FT   DISULFID     16     34        BY SIMILARITY.
FT   DISULFID     20     44        BY SIMILARITY.
FT   DISULFID     24     46        BY SIMILARITY.
FT   MOD_RES      64     64        AMIDATION.
SQ   SEQUENCE   64 AA;  6803 MW;  49A54B0B6C75392E CRC64;

Alignment Scores:
Pred. No.:                0.101          Length:         64
Score:                    77.50          Matches:        15
Percent Similarity:       51.28%         Conservative:   5
Best Local Similarity:    38.46%         Mismatches:     14
Query Match:              23.92%         Indels:         5
DB:                       1              Gaps:           2

US-09-673-274B-2 (1-169) x SCL6_LEIQH (1-64)

Qy         16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
                 :::||||||:::      :::                ||||||:::       .||||.
Db         11 AsnCysValTyrHisCysIle---------ProAspCysAspThrLeuCysLysAspAsn 27

Qy         76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCT-----AACGTGAACTGCTGGTGC 126
                ||       ||:|:|||||:|:|||                :::   |||||||||
Db         28 GlyGlyThrGlyGlyHisCysGlyPheLysLeuGlyHisGlyIleAlaCysTrpCys 46

RESULT 11
SCX4_LEIQU
ID   SCX4_LEIQU     STANDARD;      PRT;    65 AA.
AC   P01489;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Neurotoxin IV (Lqq IV) (Lqq4).
OS   Leiurus quinquestriatus quinquestriatus (Egyptian scorpion).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Leiurus.
OX   NCBI_TaxID=6885;
RN   [1]
RP   SEQUENCE.
RA   Kopeyan C., Martinez G., Rochat H.;
RT   "Primary structure of toxin IV of Leiurus quinquestriatus
RT   quinquestriatus.";
RL   FEBS Lett. 181:211-217(1985).
CC   -!- FUNCTION: Binds to sodium channels and inhibits the inactivation
CC       of the activated channels, thereby blocking neuronal transmission.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC       ALPHA-TOXIN SUBFAMILY.
DR   PIR; A01749; NTSR4L.
DR   HSSP; P45697; 1SN1.
DR   InterPro; IPR003614; Knot1.
DR   InterPro; IPR002219; Neurotoxin.
DR   InterPro; IPR002061; Scorpion_toxinL.
DR   Pfam; PF00537; toxin_3; 1.
DR   PRINTS; PR00284; TOXIN.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor;
KW   Amidation.
FT   DISULFID     13     64        BY SIMILARITY.
FT   DISULFID     17     37        BY SIMILARITY.
FT   DISULFID     23     47        BY SIMILARITY.
FT   DISULFID     27     49        BY SIMILARITY.
FT   MOD_RES      65     65        AMIDATION.
SQ   SEQUENCE   65 AA;  7202 MW;  882102EC1C2B89F6 CRC64;

Alignment Scores:
Pred. No.:                0.101          Length:         65
Score:                    77.50          Matches:        16
Percent Similarity:       57.14%         Conservative:   8
Best Local Similarity:    38.10%         Mismatches:     9
Query Match:              23.92%         Indels:         9
DB:                       1              Gaps:           4

US-09-673-274B-2 (1-169) x SCX4_LEIQU (1-65)

Qy         16 TCCTGCGTGTGG------GGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAG 69
                 :::|:||||:::      |||:::   :::|||         ||||||   ||||||
Db         12 AsnCysValTyrThrCysGlySerAsnSerTyr---------CysAsnThrGluCysThr 28

Qy         70 AGGAGGGGTTACAAGGGTGGTCACTGC---------GGTTCCTTCGCTAACGTGAACTGC 120
                 :::   |||   :::   ||||:::|||'         |||   :::   ||'      |||
Db         29 LysAsnGlyAlaGluSerGlyTyrCysGlnTrpLeuGlyLysTyrGlyAsnAla---Cys 47

Qy        121 TGGTGC 126
              |||..|
Db         48 TrpCys 49

RESULT 12
SCAT_MESMA
ID   SCAT_MESMA     STANDARD;      PRT;    85 AA.
AC   Q9UAC8;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Neurotoxin AS-1 precursor.
OS   Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Mesobuthus.
OX   NCBI_TaxID=34649;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Venom gland;
RA   Lan Z.D., Dai L., Chi C.-W.;
RT   "The gene cloning and sequencing of AS and AS-1, two novel neurotoxins
RT   from the scorpion Buthus martensi Karsch.";
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Ryanodine receptor agonist in skeletal muscle.
CC   -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC   -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
```

```
C       ALPHA-TOXIN SUBFAMILY.
C    -----------------------------------------------------------------------
C    This SWISS-PROT entry is copyright. It is produced through a collaboration
C    between the Swiss Institute of Bioinformatics and the EMBL outstation -
C    the European Bioinformatics Institute. There are no restrictions on its
C    use by non-profit institutions as long as its content is in no way
C    modified and this statement is not removed. Usage by and for commercial
C    entities requires a license agreement (See http://www.isb-sib.ch/announce/
C    or send an email to license@isb-sib.ch).
C    -----------------------------------------------------------------------
R    EMBL; AF079061; AAD47375.1; -.
R    HSSP; P01492; 1VNB.
R    InterPro; IPR003614; Knot1.
R    InterPro; IPR002061; Scorpion_toxinL.
R    Pfam; PF00537; toxin_3; 1.
R    ProDom; PD000908; Scorpion_toxinL; 1.
R    SMART; SM00505; Knot1; 1.
W    Toxin; Neurotoxin; Signal.
T    SIGNAL        1     19       POTENTIAL.
T    CHAIN        20     85       NEUROTOXIN AS-1.
T    DISULFID     31     81       BY SIMILARITY.
T    DISULFID     35     56       BY SIMILARITY.
T    DISULFID     42     63       BY SIMILARITY.
T    DISULFID     46     65       BY SIMILARITY.
Q    SEQUENCE   85 AA;  9802 MW;  E2274838439E2B95 CRC64;

lignment Scores:
red. No.:                0.119          Length:         85
core:                    77.00          Matches:        17
ercent Similarity:       53.49%         Conservative:   6
est Local Similarity:    39.53%         Mismatches:     18
uery Match:              23.77%         Indels:         2
B:                       1              Gaps:           2

S-09-673-274B-2 (1-169) x SCAT_MESMA (1-85)

y           1 GATAAGCTTATCGGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
              :::|||      |||      :::|||      :::|||      |.|   :|||::
b          26 AsnLysTyrThrGlyCysLysIleTrpCysValIleAsnAsnGluSer---CysAsnSer 44

y          61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
              |:|||||||   |||            |:|:::||       ||:      :::   |||
b          45 GluCysLysLeuArgArgGlyAsnTyrGlyTyrCys---TyrPheTrpLysLeuAlaCys 63

y         121 TGGTGCGAG 129
              :::||||||
b          64 TyrCysGlu 66

ESULT 13
IX2_LEIQU
D    SIX2_LEIQU     STANDARD;      PRT;    61 AA.
C    P19855;
T    01-FEB-1991 (Rel. 17, Created)
T    01-FEB-1991 (Rel. 17, Last sequence update)
T    28-FEB-2003 (Rel. 41, Last annotation update)
E    Insect toxin 2 (Insect toxin LqqIT2).
S    Leiurus quinquestriatus quinquestriatus (Egyptian scorpion).
C    Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
C    Buthoidea; Buthidae; Leiurus.
X    NCBI_TaxID=6885;
N    [1]
P    SEQUENCE.
X    MEDLINE=90184494; PubMed=2311768;
A    Kopeyan C., Mansuelle P., Sampieri F., Brando T., Bahraoui E.M.,
A    Rochat H., Granier C.;
T    "Primary structure of scorpion anti-insect toxins isolated from the
T    venom of Leiurus quinquestriatus quinquestriatus.";
L    FEBS Lett. 261:423-426(1990).
N    [2]
P    SEQUENCE.
X    MEDLINE=93160530; PubMed=8431601;
A    Zlotkin E., Gurevitz M., Fowler E., Adams M.E.;
T    "Depressant insect selective neurotoxins from scorpion venom:
```

```
RT   chemistry, action, and gene cloning.";
RL   Arch. Insect Biochem. Physiol. 22:55-73(1993).
RN   [3]
RP   ACTION.
RA   Zlotkin E., Fowler E., Eitan M., Moyer M., Adams M.E.;
RT   "On the chemistry and action of the depressant insect toxins.";
RL   Toxicon 28:170-170(1990).
CC   -!- FUNCTION: Binds to sodium channels and inhibits the inactivation
CC       of the activated channels, thereby blocking neuronal transmission.
CC       This toxin induces a slow, depressant, flaccid paralysis on fly
CC       larvae. It is active only on insects.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC       ALPHA-TOXIN SUBFAMILY.
DR   PIR; B34123; B34123.
DR   HSSP; P01494; 2SN3.
DR   InterPro; IPR003614; Knot1.
DR   InterPro; IPR002061; Scorpion_toxinL.
DR   Pfam; PF00537; toxin_3; 1.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor.
FT   DISULFID     10     60       BY SIMILARITY.
FT   DISULFID     14     35       BY SIMILARITY.
FT   DISULFID     21     42       BY SIMILARITY.
FT   DISULFID     25     44       BY SIMILARITY.
SQ   SEQUENCE   61 AA;  6860 MW;  80F01771C7DCFAF6 CRC64;

Alignment Scores:
Pred. No.:              0.134          Length:         61
Score:                  76.50          Matches:        15
Percent Similarity:     55.26%         Conservative:   6
Best Local Similarity:  39.47%         Mismatches:     12
Query Match:            23.61%         Indels:         5
DB:                     1              Gaps:           2

US-09-673-274B-2 (1-169) x SIX2_LEIQU (1-61)

Qy         16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
              :|||: ::::::|||                      |||||   ||:| ||||
Db         13 SerCysLeuPheGly----- ------AsnGluGlyCysAsnLysGluCysLysSerTyr 28

Qy         76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
              |||         |||:::|||   ::::::      :::   '|||||:|.|||'
Db         29 GlyGlySerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 45

RESULT 14
SCXE_BUTEU
ID   SCXE_BUTEU     STANDARD;      PRT;    66 AA.
AC   P09982;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-MAR-1989 (Rel. 10, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Neurotoxin M14.
OS   Buthus eupeus (Lesser Asian scorpion) (Mesobuthus eupeus).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Mesobuthus.
OX   NCBI_TaxID=34648;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Venom;
RX   MEDLINE=85046632; PubMed=6497916;
RA   Volkova T.M., Garsia A.F., Telezhinskaya I.N., Potapenko N.A.,
RA   Grishin E.V.;
RT   "Amino acid sequence of 2 neurotoxins from the scorpion Buthus eupeus
RT   venom.";
RL   Bioorg. Khim. 10:979-982(1984).
CC   -!- FUNCTION: Binds to sodium channels and inhibits the inactivation
CC       of the activated channels, thereby blocking neuronal transmission.
CC       Has paralytic activity in mice.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom gland.
```

```
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC       ALPHA-TOXIN SUBFAMILY.
DR   PIR; JT0220; NTSR4E.
DR   HSSP; P45697; 1SN1.
DR   InterPro; IPR003614; Knot1.
DR   InterPro; IPR001219; Neurotoxin.
DR   InterPro; IPR002061; Scorpion_toxinL.
DR   Pfam; PF00537; toxin_3; 1.
DR   PRINTS; PR00284; TOXIN.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor.
FT   DISULFID     12     65       BY SIMILARITY.
FT   DISULFID     16     36       BY SIMILARITY.
FT   DISULFID     22     46       BY SIMILARITY.
FT   DISULFID     26     48       BY SIMILARITY.
SQ   SEQUENCE   66 AA;  7453 MW;  B5E0DB42237EF2E8 CRC64;

Alignment Scores:
Pred. No.:                 0.155          Length:        66
Score:                     76.00          Matches:       13
Percent Similarity:        58.62%         Conservative:  4
Best Local Similarity:     44.83%         Mismatches:    8
Query Match:               23.46%         Indels:        4
DB:                        1              Gaps:          2

US-09-673-274B-2 (1-169) x SCXE_BUTEU (1-66)

Qy          52 TGCAACGGTGAGTGCAAGAGGAGGGTTACAAGGGTGGTCACTGC---------GGTTCC 102
               |: :::   |:||||||| :::    .||          |..   :::|||          | :
Db          22 CysAspSerGluCysLysLysAsnGlyAlaAspGlySerTyrCysGlnTrpLeuGlyArg 41

Qy         103 TTCGCTAACGTGAACTGCTGGTGCGAG 129
                 :     ||          | | ||||::::
Db          42 PheGlyAsnAla---CysTrpCysLys 49

RESULT 15
SCL3_LEIQH
ID   SCL3_LEIQH     STANDARD;      PRT;    67 AA.
AC   P56678;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Alpha-like neurotoxin Lqh III (Lqh3).
OS   Leiurus quinquestriatus hebraeus (Yellow scorpion).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Leiurus.
OX   NCBI_TaxID=6884;
RN   [1]
RP   SEQUENCE, PHARMACOLOGICAL CHARACTERIZATION, AND BIOASSAY.
RC   TISSUE=Venom;
RX   MEDLINE=98353091; PubMed=9690781;
RA   Sautiere P., Cestele S., Kopeyan C., Martinage A., Drobecq H.,
RA   Doljansky Y., Gordon D.;
RT   "New toxins acting on sodium channels from the scorpion Leiurus
RT   quinquestriatus hebraeus suggest a clue to mammalian vs insect
RT   selectivity.";
RL   Toxicon 36:1141-1154(1998).
RN   [2]
RP   STRUCTURE BY NMR.
RC   TISSUE=Venom;
RX   MEDLINE=99141414; PubMed=9917409;
RA   Krimm I., Gilles N., Sautiere P., Stankiewicz M., Pelhate M.,
RA   Gordon D., Lancelin J.-M.;
RT   "NMR structures and activity of a novel alpha-like toxin from the
RT   scorpion Leiurus quinquestriatus hebraeus.";
RL   J. Mol. Biol. 285:1749-1763(1999).
RN   [3]
RP   STRUCTURE BY NMR.
RC   TISSUE=Venom;
RA   Krimm I., Trivelli X., Lancelin J.-M.;
RT   "A cis-trans isomerism of a non-prolyl peptide bond in LQH III alpha-
RT   like scorpion toxin revealed by solution NMR.";
```

```
RL   Submitted (JUL-2000) to the PDB data bank.
CC   -!- FUNCTION: Binds to sodium channels and inhibits the inactivation
CC       of the activated channels, thereby blocking neuronal transmission.
CC       Active on both insects and mammals but competes for alpha-toxins
CC       binding only on cockroach sodium channels.
CC   -!- SUBUNIT: Monomer.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC   -!- MISCELLANEOUS: LD(50) is 50 mg/kg by intracerebroventricular
CC       injection into mice.
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC       ALPHA-TOXIN SUBFAMILY. ALPHA-LIKE TOXIN SUBCLASS.
DR   PDB; 1BMR; 16-FEB-99.
DR   PDB; 1FH3; 08-NOV-00.
DR   InterPro; IPR003614; Knot1.
DR   InterPro; IPR002061; Scorpion_toxinL.
DR   Pfam; PF00537; toxin_3; 1.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor;
KW   Amidation; 3D-structure.
FT   DISULFID     12     65
FT   DISULFID     16     37
FT   DISULFID     23     47
FT   DISULFID     27     49
FT   MOD_RES      67     67       AMIDATION.
FT   STRAND        2      6
FT   TURN          9     11
FT   HELIX        21     30
FT   TURN         31     31
FT   STRAND       34     40
FT   TURN         41     43
FT   STRAND       44     52
FT   TURN         54     55
SQ   SEQUENCE   67 AA;  7057 MW;  19FE8EF96154328F CRC64;

Alignment Scores:
Pred. No.:                 0.155          Length:        67
Score:                     76.00          Matches:       14
Percent Similarity:        51.28%         Conservative:  6
Best Local Similarity:     35.90%         Mismatches:    17
Query Match:               23.46%         Indels:        2
DB:                        1              Gaps:          1

US-09-673-274B-2 (1-169) x SCL3_LEIQH (1-67)

Qy          16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
               :::||| ||:::                     :::|||    |||:::          ||||||    :::
Db          11 AsnCysValTyrHisCysPheProGlySerSerGlyCysAspThrLeuCysLysGluLys 30

Qy          76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCT------AACGTGAACTGCTGGTGC 126
               :||            |||||||||||:|                    :::    |||||||||
Db          31 GlyGlyThrSerGlyHisCysGlyPheLysValGlyHisGlyLeuAlaCysTrpCys 49


Search completed: October 16, 2003, 17:07:23
Job time : 12.2292 secs
```

```
                          GenCore version 5.1.6
                 Copyright (c) 1993 - 2003  Compugen Ltd.

M nucleic - protein search, using frame_plus_n2p model

un on:          October 16, 2003, 17:06:37 ; Search time 49.9701 Seconds
                                              (without alignments)
                                              1745.479 Million cell updates/sec

itle:           US-09-673-274B-2
erfect score:   324
equence:        1 gataagcttatcggttcctg..........gaacgtgtcgacggatccgg 169

coring table:   BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

earched:        830525 seqs, 258052604 residues

otal number of hits satisfying chosen parameters:      1661050

inimum DB seq length: 0
aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

ommand line parameters:
MODEL=frame+_n2p.model -DEV=xlp
Q=/cgn2_1/USPTO_spool/US09673274/runat_16102003_170627_2545/app_query.fasta_1.654
DB=SPTREMBL_23 -QFMT=fastan -SUFFIX=rspt -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
USER=US09673274_@CGN_1_1_216_@runat_16102003_170627_2545 -NCPU=6 -ICPU=3
NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

atabase :       SPTREMBL_23:*
                1:  sp_archea:*
                2:  sp_bacteria:*
                3:  sp_fungi:*
                4:  sp_human:*
                5:  sp_invertebrate:*
                6:  sp_mammal:*
                7:  sp_mhc:*
                8:  sp_organelle:*
                9:  sp_phage:*
                10:  sp_plant:*
                11:  sp_rodent:*
                12:  sp_virus:*
                13:  sp_vertebrate:*
                14:  sp_unclassified:*
                15:  sp_rvirus:*
                16:  sp_bacteriap:*
                17:  sp_archeap:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                   SUMMARIES
                  %
esult             Query
 No.    Score  Match Length DB  ID                     Description
------------------------------------------------------------------------------
   1     84.5   26.1     65  5  Q17231                 Q17231 mesobuthus
```

```
       2    83.5   25.8     65  5  Q17230                     Q17230 mesobuthus
       3    83.5   25.8     85  5  Q8I0K7                     Q8i0k7 mesobuthus
       4    81.5   25.2     85  5  Q95WX6                     Q95wx6 mesobuthus
       5    81.5   25.2     85  5  Q9XY87                     Q9xy87 mesobuthus
       6    80.5   24.8     85  5  Q8T3T0                     Q8t3t0 mesobuthus
       7      80   24.7     91 10   P82761                     P82761 arabidopsis
       8    78.5   24.2     69  5  Q9VZQ5                     Q9vzq5 drosophila
       9    76.5   23.6     82  5  Q8MVS7                     Q8mvs7 leiurus qui
      10    76.5   23.6     85  5  Q9Y1U3                     Q9y1u3 mesobuthus
 c    11      75   24.8    173  5  Q9N670                     Q9n670 leishmania
      12    74.5   23.0     69  5  Q9VZR2                     Q9vzr2 drosophila
      13    74.5   23.0     84  5  Q8WRY4                     Q8wry4 centruroide
      14      74   22.8     85  5  Q9GUA7                     Q9gua7 mesobuthus
      15    73.5   22.7    106  5  Q9VZR3                     Q9vzr3 drosophila
      16      72   22.2     85  5  Q95P69                     Q95p69 mesobuthus
      17    71.5   22.1     73 10   Q9FPG1                     Q9fpg1 phaseolus a
      18    71.5   22.1     82  5  Q9U6U0                     Q9u6u0 mytilus gal
      19    71.5   22.1     82  5  Q8MVS6                     Q8mvs6 leiurus qui
      20    71.5   22.1     82  5  Q8MVS8                     Q8mvs8 leiurus qui
      21      71   21.9     81  5  Q9Y0A9                     Q9y0a9 mytilus gal
      22      71   21.9   1045  5  Q8T3A6                     Q8t3a6 caenorhabdi
      23      71   21.9   1070  5  Q8T3A7                     Q8t3a7 caenorhabdi
      24      71   21.9   1111  5  Q9XWD6                     Q9xwd6 caenorhabdi
      25    70.5   21.8     71  5  Q8IRD6                     Q8ird6 drosophila
      26    70.5   21.8     72  5  Q9VZQ4                     Q9vzq4 drosophila
      27    68.5   21.1     85  5  Q9Y0B8                     Q9y0b8 mesobuthus
      28      68   21.0     77 10   Q39403                     Q39403 brassica ca
 c    29      67   22.1    469  5  Q9U1K6                     Q9u1k6 drosophila
 c    30      67   22.1    560  5  Q9W526                     Q9w526 drosophila
 c    31    66.5   21.9    873 16   Q9RX36                     Q9rx36 deinococcus
 c    32    66.5   21.9   1275  4  Q9UQ36                     Q9uq36 homo sapien
      33    66.5   20.5   1405 11   Q8VHS2                     Q8vhs2 mus musculu
 c    34    66.5   21.9   1783  4  O15038                     O15038 homo sapien
 c    35    66.5   21.9   1791  4  O60382                     O60382 homo sapien
 c    36    66.5   21.9   2296  4  Q9UHA8                     Q9uha8 homo sapien
 c    37    66.5   21.9   2752  4  Q9UQ35                     Q9uq35 homo sapien
      38      66   20.4     55  5  Q9TWL1                     Q9twl1 centruroide
      39      66   20.4   3704  5  P91904                     P91904 caenorhabdi
      40    65.5   20.2     64  5  O76963                     O76963 orthochirus
      41    65.5   20.2    428 12   Q99H52                     Q99h52 nectria rad
      42      65   20.1     85  5  Q9BLM4                     Q9blm4 androctonus
 c    43      65   21.5    206 12   Q9WHF2                     Q9whf2 agrotis seg
 c    44      65   21.5    628  5  Q9VSB3                     Q9vsb3 drosophila
      45    64.5   19.9     85  5  Q26292                     Q26292 leiurus qui


                                  ALIGNMENTS

RESULT 1
Q17231
ID   Q17231      PRELIMINARY;      PRT;    65 AA.
AC   Q17231;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Insect toxin 2.
GN   IT2.
OS   Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Mesobuthus.
OX   NCBI_TaxID=34649;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Telson;
RA   Zhu X., Zhang T., Zhu Y.;
RT   "Cloning and sequencing of two depressant insect selective neurotoxin
RT   cDNAs from Buthus martensii Karsch.";
RL   Kexue Tongbao 41:1387-1391(1996).
CC   -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
DR   EMBL; X92077; CAA63059.1; -.
DR   HSSP; P01494; 2SN3.
DR   InterPro; IPR003614; Knot1.
```

```
DR   InterPro; IPR001219; Neurotoxin.
DR   InterPro; IPR002061; Scorpion_toxinL.
DR   Pfam; PF00537; toxin_3; 1.
DR   PRINTS; PR00284; TOXIN.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
SQ   SEQUENCE   65 AA;  7285 MW;  F13FCD7153855626 CRC64;

Alignment Scores:
Pred. No.:              0.0313       Length:        65
Score:                  84.50        Matches:       15
Percent Similarity:     55.26%       Conservative:  6
Best Local Similarity:  39.47%       Mismatches:    12
Query Match:            26.08%       Indels:        5
DB:                     5            Gaps:          2

US-09-673-274B-2 (1-169) x Q17231 (1-65)

Qy          16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
               ||||||:::||||||                     ||||||   ||||||:::
Db          14 SerCysLeuTrpGly------------AsnGluGlyCysAsnLysGluCysArgAlaTyr 29

Qy          76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
               |||         |||:::|||   ::::::      :::   ||||||||||||
Db          30 GlyAlaSerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 46

RESULT 2
Q17230
ID   Q17230      PRELIMINARY;      PRT;    65 AA.
AC   Q17230;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Insect-toxin 3.
GN   IT3.
OS   Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Mesobuthus.
OX   NCBI_TaxID=34649;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Telson;
RA   Zhu X., Zhang T., Zhu Y.;
RT   "Cloning and sequencing of two depressant insect selective neurotoxin
RT   cDNAs from Buthus martensii Karsch.";
RL   Kexue Tongbao 41:1387-1391(1996).
CC   -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
DR   EMBL; X92846; CAA63431.1; -.
DR   HSSP; P01491; 1B3C.
DR   InterPro; IPR003614; Knot1.
DR   InterPro; IPR001219; Neurotoxin.
DR   InterPro; IPR002061; Scorpion_toxinL.
DR   Pfam; PF00537; toxin_3; 1.
DR   PRINTS; PR00284; TOXIN.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
SQ   SEQUENCE   65 AA;  7317 MW;  ECDCC391B733EC0A CRC64;

Alignment Scores:
Pred. No.:              0.0422       Length:        65
Score:                  83.50        Matches:       16
Percent Similarity:     55.26%       Conservative:  5
Best Local Similarity:  42.11%       Mismatches:    12
Query Match:            25.77%       Indels:        5
DB:                     5            Gaps:          2

US-09-673-274B-2 (1-169) x Q17230 (1-65)

Qy          16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
               | |||||:::||||||                    ||||||   |||||||||
Db          14 SerCysLeuTrpGly------------AsnGluGlyCysAsnLysGluCysLysGlyPhe 29
```

```
Qy          76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
               |||         |||:::|||   ::::::      :::   ||||||||||||
Db          30 GlyAlaTyrTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 46

RESULT 3
Q8I0K7
ID   Q8I0K7      PRELIMINARY;      PRT;    85 AA.
AC   Q8I0K7;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Depressant scorpion toxin KIM2 precursor.
GN   KIM2.
OS   Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Mesobuthus.
OX   NCBI_TaxID=34649;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22217688; PubMed=12230558;
RA   Peng F., Zeng X.C., He X.H., Pu J., Li W.X., Zhu Z.H., Liu H.;
RT   "Molecular cloning and functional expression of a gene encoding an
RT   antiarrhythmia peptide derived from the scorpion toxin.";
RL   Eur. J. Biochem. 269:4468-4475(2002).
DR   EMBL; AF459791; AAN59782.1; -.
DR   EMBL; AF459792; AAN59783.1; -.
SQ   SEQUENCE   85 AA;  9425 MW;  5076039B9966AA6F CRC64;

Alignment Scores:
Pred. No.:              0.043        Length:        85
Score:                  83.50        Matches:       16
Percent Similarity:     55.26%       Conservative:  5
Best Local Similarity:  42.11%       Mismatches:    12
Query Match:            25.77%       Indels:        5
DB:                     5            Gaps:          2

US-09-673-274B-2 (1-169) x Q8I0K7 (1-85)

Qy          16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
               ||||||:::||||||                     ||||||   |||||||||
Db          34 SerCysLeuTrpGly------------AsnGluGlyCysAsnLysGluCysLysGlyPhe 49

Qy          76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
               |||         |||:::|||   ::::::      :::   ||||||||||||
Db          50 GlyAlaTyrTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 66

RESULT 4
Q95WX6
ID   Q95WX6      PRELIMINARY;      PRT;    85 AA.
AC   Q95WX6;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Depressant insect neurotoxin BmK ITb precursor.
OS   Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Mesobuthus.
OX   NCBI_TaxID=34649;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Wang C., Ling M., Wang D., Chi C.;
RT   "The gene for depressant insect toxin BmK ITb from Chinese scorpion,
RT   Buthus martensii Karsch.";
RL   Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
DR   EMBL; AF272777; AAF77063.2; -.
DR   InterPro; IPR003614; Knot1.
DR   InterPro; IPR002061; Scorpion_toxinL.
DR   Pfam; PF00537; toxin_3; 1.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Neurotoxin; Signal.
```

```
T    SIGNAL        1     21        POTENTIAL.
T    CHAIN        22     82        DEPRESSANT INSECT NEUROTOXIN BMK ITB.
Q    SEQUENCE   85 AA;  9330 MW;  80CDD3914956DE8C CRC64;

lignment Scores:
red. No.:                0.0783          Length:          85
core:                    81.50           Matches:         15
ercent Similarity:       55.26%          Conservative:    6
est Local Similarity:    39.47%          Mismatches:      12
uery Match:              25.15%          Indels:          5
B:                       5               Gaps:            2 .

S-09-673-274B-2 (1-169) x Q95WX6 (1-85)

y          16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
              ||||||:::||||||                     ||||||   |||:|||:.
b          34 SerCysLeuTrpGly------------AsnGluGlyCysAsnLysGluCysLysAlaPhe 49

y          76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
              |||         |||:::|||   ::::::      :::   |||||||||:::
b          50 GlyAlaTyrTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGln 66

ESULT 5
9XY87
D    Q9XY87       PRELIMINARY;      PRT;    85 AA.
C    Q9XY87;
T    01-NOV-1999 (TrEMBLrel. 12, Created)
T    01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
T    01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
E    Neurotoxin AEP precursor.
S    Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).
C    Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
C    Buthoidea; Buthidae; Mesobuthus.
X    NCBI_TaxID=34649;
N    [1]
P    SEQUENCE FROM N.A.
A    Ling M.H., Wang C.G., Wang D.C., Chi C.W.;
L    Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
C    -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
C    -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
R    EMBL; AF064821; AAD31592.1; -.
R    HSSP; P01494; 2SN3.
R    InterPro; IPR003614; Knot1.
R    InterPro; IPR002061; Scorpion_toxinL.
R    Pfam; PF00537; toxin_3; 1.
R    ProDom; PD000908; Scorpion_toxinL; 1.
R    SMART; SM00505; Knot1; 1.
W    Neurotoxin; Signal.
T    SIGNAL        1     21        POTENTIAL.
T    CHAIN        22     85        NEUROTOXIN AEP.
Q    SEQUENCE   85 AA;  9312 MW;  CCE46711BAF21DAC CRC64;

lignment Scores:
red. No.:                0.0783          Length:          85
core:                    81.50           Matches:         14
ercent Similarity:       55.26%          Conservative:    7
est Local Similarity:    36.84%          Mismatches:      12
uery Match:              25.15%          Indels:          5
B:                       5               Gaps:            2

S-09-673-274B-2 (1-169) x Q9XY87 (1-85)

y          16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
              ||||||:::||||||                     |:|||:   ||||||:::
b          34 SerCysLeuTrpGly------------AsnGluGlyCysAsnLysGluCysArgAlaTyr 49

y          76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
              |||         |||:::|||   ::::::      :::   |||||||||:::
b          50 GlyAlaSerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGln 66

ESULT 6
8T3T0
D    Q8T3T0       PRELIMINARY;      PRT;    85 AA.
```

```
AC   Q8T3T0;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Depressant insect toxin BmK ITa1 precursor.
OS   Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Mesobuthus.
OX   NCBI_TaxID=34649;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Liu Z., Chi C., Wu X.;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
DR   EMBL; AY090782; AAM09097.1; -.
DR   InterPro; IPR003614; Knot1.
DR   InterPro; IPR002061; Scorpion_toxinL.
DR   Pfam; PF00537; toxin_3; 1.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Signal.
FT   SIGNAL        1     21        POTENTIAL.
SQ   SEQUENCE   85 AA;  9214 MW;  6CF5670B6032001A CRC64;

Alignment Scores:
Pred. No.:               0.106           Length:          85
Score:                   80.50           Matches:         15
Percent Similarity:      52.63%          Conservative:    5
Best Local Similarity:   39.47%          Mismatches:      13
Query Match:             24.85%          Indels:          5
DB:                      5               Gaps:            2

US-09-673-274B-2 (1-169) x Q8T3T0 (1-85)

Qy         16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
              | |||:::||: ||                      ||||||   |||||:.
Db         34 SerCysLeuTrpGly------------AsnGluGlyCysAsnLysGluCysGlyAlaTyr 49

Qy         76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
              |.|         ||::::||    ::::::      :::   :|||||:|.|||
Db         50 GlyAlaSerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 66

RESULT 7
P82761
ID   P82761       PRELIMINARY;      PRT;    91 AA.
AC   P82761;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Hypothetical protein LCR46 precursor.
GN   LCR46.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=99397451; PubMed=10470850;
RA   Kaneko T., Katoh T., Sato S., Nakamura Y., Asamizu E., Kotani H.,
RA   Miyajima N., Tabata S.;
RT   "Structural analysis of Arabidopsis thaliana chromosome 5. IX.
RT   Sequence features of the regions of 1,011,550 bp covered by seventeen
RT   P1 and TAC clones.";
RL   DNA Res. 6:183-195(1999).
RN   [2]
RP   IDENTIFICATION.
RX   PubMed=11437247;
RA   Vanoosthuyse V., Miege C., Dumas C., Cock J.M.;
RT   "Two large Arabidopsis thaliana gene families are homologous to the
RT   Brassica gene superfamily that encodes pollen coat proteins and the
```

```
RT   male component of the self-incompatibility response.";
RL   Plant Mol. Biol. 46:17-34(2001).
DR   EMBL; AB017061; -; NOT_ANNOTATED_CDS.
KW   Hypothetical protein; Signal.
FT   SIGNAL        1     27       POTENTIAL.
FT   CHAIN        28     91       HYPOTHETICAL PROTEIN LCR46.
SQ   SEQUENCE   91 AA;  9503 MW;  10BDEAAE80969F55 CRC64;

Alignment Scores:
Pred. No.:                0.123         Length:         91
Score:                    80.00         Matches:        14
Percent Similarity:       58.06%        Conservative:   4
Best Local Similarity:    45.16%        Mismatches:     13
Query Match:              24.69%        Indels:         0
DB:                       10            Gaps:           0

US-09-673-274B-2 (1-169) x P82761 (1-91)

Qy          16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
               :::| |:::   :||               || :::   |||:     ||||:
Db          31 ThrCysLeuProGlyGluCysThrAsnProSerGluCysAsnAlaAlaCysLysSerAsn 50

Qy          76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCT 108
               |:::|||||:|:|||   :|   ||:  :::
Db          51 GlyTyrLysGlyGlyAlaCysValSerMetSer 61

RESULT 8
Q9VZQ5
ID   Q9VZQ5       PRELIMINARY;      PRT;    69 AA.
AC   Q9VZQ5;
DT   01-MAY 2000 (TrEMBLrel. 13, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01 MAR 2003 (TrEMBLrel. 23, Last annotation update)
DE   CG32274 protein.
GN   CG32268 OR CG10815 OR CG11520.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA   Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA   Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA   Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA   Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA   Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA   Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA   Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA   McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA   Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA   Phouanenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA   Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA   Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT   "Sequencing of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA   Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA   Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA   Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA   Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA   Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA   Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT   "Annotation of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RA   FlyBase;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003477; AAF47764.2; -.
DR   HSSP; P41964; 1MYN.
DR   FlyBase; FBgn0052268; CG32268.
DR   InterPro; IPR002118; Gamma-thionin.
DR   ProDom; PD002594; Gamma-thionin; 1.
SQ   SEQUENCE   69 AA;  7720 MW;  87CC05D254A11AC4 CRC64;

Alignment Scores:
Pred. No.:                0.19          Length:         69
Score:                    78.50         Matches:        16
Percent Similarity:       51.28%        Conservative:   4
Best Local Similarity:    41.03%        Mismatches:     14
Query Match:              24.23%        Indels:         5
DB:                       5             Gaps:           2

US-09-673-274B-2 (1-169) x Q9VZQ5 (1-69)

Qy          13 GGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGG 72
               |||:||||:         ||||||   :::         |||         |||:::|||
Db          34 GlySerCys---------AlaValTrpHisArgLysLysCysValAspIleCysGlnArg 50

Qy          73 AGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
                  |||         ||:|||||:            ::::::   |||||||||||
Db          51 GluGlyArgThrSerGlyHisCysSer------ProSerLeuLysCysTrpCysGlu 67

RESULT 9
Q8MVS7
ID   Q8MVS7       PRELIMINARY;      PRT;    82 AA.
```

```
AC   Q8MVS7;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Depressant insect toxin 2.
OS   Leiurus quinquestriatus quinquestriatus (Egyptian scorpion).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Leiurus.
OX   NCBI_TaxID=6885;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Zaki T.I., Maruniak J.E.;
RT   "Three polymorphic genes encoding LqqIT2 from the Egyptian scorpion
RT   Leiurus quinquestriatus quinquestriatus.";
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
DR   EMBL; AF474984; AAM74028.1; -.
DR   InterPro; IPR003614; Knot1.
DR   InterPro; IPR002061; Scorpion_toxinL.
DR   Pfam; PF00537; toxin_3; 1.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
SQ   SEQUENCE   82 AA;  9099 MW;  F13D53B18CDECBF0 CRC64;

Alignment Scores:
Pred. No.:              0.35          Length:        82
Score:                  76.50         Matches:       15
Percent Similarity:     55.26%        Conservative:  6
Best Local Similarity:  39.47%        Mismatches:    12
Query Match:            23.61%        Indels:        5
DB:                     5             Gaps:          2

US-09-673-274B-2 (1-169) x Q8MVS7 (1-82)

Qy         16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
              ||||||::::::|||                     ||:|||   ||| ||||:
Db         34 SerCysLeuPheGly------------AsnGluGlyCysAsnLysGluCysLysSerTyr 49

Qy         76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
              |||         |||:::|||   ::::::      :::   ||||||||||||
Db         50 GlyGlySerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 66

RESULT 10
Q9Y1U3
ID   Q9Y1U3      PRELIMINARY;      PRT;    85 AA.
AC   Q9Y1U3;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Depressant insect neurotoxin precursor.
GN   ITC.
OS   Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Mesobuthus.
OX   NCBI_TaxID=34649;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ling M., Wang C., Wang D., Chi C.;
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
DR   EMBL; AF073899; AAD41648.2; -.
DR   HSSP; P01491; 1B3C.
DR   InterPro; IPR003614; Knot1.
DR   InterPro; IPR002061; Scorpion_toxinL.
DR   Pfam; PF00537; toxin_3; 1.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Neurotoxin; Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22     82       DEPRESSANT INSECT NEUROTOXIN.
SQ   SEQUENCE   85 AA;  9480 MW;  18E02CD791F3B5EB CRC64;

Alignment Scores:
Pred. No.:              0.351         Length:        85
Score:                  76.50         Matches:       14
Percent Similarity:     55.26%        Conservative:  7
Best Local Similarity:  36.84%        Mismatches:    12
Query Match:            23.61%        Indels:        5
DB:                     5             Gaps:          2

US-09-673-274B-2 (1-169) x Q9Y1U3 (1-85)

Qy         16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
              ||:||::::|||:|                      |||:::      |||::::::
Db         34 SerCysLeuTrpGly------------AsnAspPheCysAspLysValCysLysLysSer 49

Qy         76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
              ||:         |:|:::|:    ::::::      :::   |||:|:|||:::
Db         50 GlyGlySerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 66

RESULT 11
Q9N670
ID   Q9N670      PRELIMINARY;      PRT;   173 AA.
AC   Q9N670;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Ribosomal protein L38.
GN   RPL38.
OS   Leishmania major.
OC   Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX   NCBI_TaxID=5664;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FRIEDLIN;
RA   Myler P.J.;
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FRIEDLIN;
RA   Myler P.J., Sisk E., Hixson G., Kiser P., Rickel E., Hassebrock M.,
RA   Cawthra J., Marsolini F., Sunkin S., Stuart K.D.;
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE L38E FAMILY OF RIBOSOMAL PROTEINS.
DR   EMBL; AC005802; AAF31045.1; -.
DR   EMBL; AC005893; AAF31039.1; -.
DR   InterPro; IPR002675; Ribosomal_L38e.
DR   Pfam; PF01781; Ribosomal_L38e; 1.
DR   ProDom; PD010361; Ribosomal_L38e; 1.
KW   Ribonucleoprotein; Ribosomal protein.
SQ   SEQUENCE   173 AA;  19105 MW;  CE1433DAE9E81F2E CRC64;

Alignment Scores:
Pred. No.:              0.577         Length:        173
Score:                  75.00         Matches:       22
Percent Similarity:     41.79%        Conservative:  6
Best Local Similarity:  32.84%        Mismatches:    19
Query Match:            24.75%        Indels:        20
DB:                     5             Gaps:          3

US-09-673-274B-2 (1-169) x Q9N670 (1-173)

Qy        166 GATCCGTCGACACGTTCGCCTCGCCGAGCTCTCAAGTCTCGC------------------ 125
              |||   |||:||||  |||::|         |:|   | ||||
Db         13 AspAlaSerThrArgSerProThrHisThrLeuProSerArgCysHisLeuPheLeuSer 32

Qy        124 ------ACCAGCAGTTCACGTTAGCGAAGGAACCGCAGTGACCACCCTTGTAACCCCTCC 71
                       |||::::::      ||||||:::|||:::      |||      |:::::
Db         33 LeuSerLeuSerThrAlaAlaValArgArgSerArgAlaSerProProLeuAlaProThr 52

Qy         70 TCTTGCACTCACCGTTGC---------------AATCGGAAGTGTAGTTCA--------- 35
                    ||:|||   |||                   ||   |||
Db         53 ArgArgThrHisGluCysLeuSerAlaSerHisThrArgThrCysArgGlnSerAlaSer 72
```

```
Qy          34 ------------CAGCACCCC 26
                           :::||||.|
Db          73 PheSerProLeuGluHisPro 79

RESULT 12
Q9VZR2
ID   Q9VZR2      PRELIMINARY;      PRT;    69 AA.
AC   Q9VZR2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   CG10812 protein (GH09576p).
GN   CG10812.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA   George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA   Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA   Patel S., Phouanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA   Celniker S.;
RL   Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003477; AAF47757.1; -.
DR   EMBL; AY118762; AAM50622.1; -.
DR   HSSP; P41964; 1MYN.
DR   FlyBase; FBgn0035434; CG10812.
DR   InterPro; IPR002118; Gamma-thionin.
DR   ProDom; PD002594; Gamma-thionin; 1.
SQ   SEQUENCE   69 AA;  7656 MW;  72555FD8B1DCC6F7 CRC64;

Alignment Scores:
Pred. No.:              0.63          Length:        69
Score:                  74.50         Matches:       15
Percent Similarity:     42.50%        Conservative:  2
Best Local Similarity:  37.50%        Mismatches:    16
Query Match:            22.99%        Indels:        7
DB:                     5             Gaps:          3

US-09-673-274B-2 (1-169) x Q9VZR2 (1-69)

Qy          13 GGTTCCTGC---GTGTGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAG 69
               |||   |||   ||||||                           |||         ||||||
Db          34 GlyProCysAlaValTrpAsp------------AsnGluThrCysArgArgValCysLys 49

Qy          70 AGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
                     |:|         |||||||||             ::::::      ||||||||||||
Db          50 GluGluGlyArgSerSerGlyHisCysSer------ProSerLeuLysCysTrpCysGlu 67

RESULT 13
Q8WRY4
ID   Q8WRY4      PRELIMINARY;      PRT;    84 AA.
AC   Q8WRY4;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Na+-channel modifying toxin precursor.
GN   CLL9.
OS   Centruroides limpidus limpidus (Mexican scorpion).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Centruroides.
OX   NCBI_TaxID=29941;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Corona M., Coronas F., Merino E., Becerril B., Antunez S.R.,
RA   Garcia D.E., Possani L.D.;
RT   "Cll9, a novel class of peptide found in scorpion venom which
RT   particularly affects Na+-channels of rat peripheral ganglia
RT   preparations.";
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
DR   EMBL; AF387171; AAL57291.1; -.
DR   InterPro; IPR003614; Knot1.
DR   InterPro; IPR002061; Scorpion_toxinL.
DR   Pfam; PF00537; toxin_3; 1.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Signal.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20     82       NA+-CHANNEL MODIFYING TOXIN.
SQ   SEQUENCE   84 AA;  9198 MW;  21018F12E2ECC26E CRC64;

Alignment Scores:
Pred. No.:              0.638         Length:        84
Score:                  74.50         Matches:       14
Percent Similarity:     48.72%        Conservative:  5
Best Local Similarity:  35.90%        Mismatches:    15
Query Match:            22.99%        Indels:        5
DB:                     5             Gaps:          2

US-09-673-274B-2 (1-169) x Q8WRY4 (1-84)

Qy          19 TGCGTGTGGGGTGCTGTGAACTACACTTCCGAT------TGCAACGGTGAGTGCAAGAGG 72
               |||            ::::::   |||   |||      |||:::      ||||||:::
Db          31 CysThrLeuAlaCysIleAspLysThrGlyAspLysAsnCysAspArgAsnCysLysLys 50

Qy          73 AGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
                  |||         |||||||||         |||         |||||||||:::
```

```
b         51 GluGlyGlySerPheGlyHisCysSerTyrSerAla---------CysTrpCysLys 66

ESULT 14
9GUA7
D    Q9GUA7        PRELIMINARY;      PRT;    85 AA.
C    Q9GUA7;
T    01-MAR-2001 (TrEMBLrel. 16, Created)
T    01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
T    01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
E    Hypothetical 9.4 kDa protein.
S    Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).
C    Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
C    Buthoidea; Buthidae; Mesobuthus.
X    NCBI_TaxID=34649;
N    [1]
P    SEQUENCE FROM N.A.
C    TISSUE=Venom gland;
A    Li W.-X., Zeng X.-C., Zu S.-Y.;
T    "A new cDNA sequence encoding a toxin from Buthus martensii Karsch.";
L    Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
C    -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
C    -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
R    EMBL; AF150009; AAG09657.1; -.
R    HSSP; P17728; 1LQI.
R    InterPro; IPR003614; Knot1.
R    InterPro; IPR001219; Neurotoxin.
R    InterPro; IPR002061; Scorpion_toxinL.
R    Pfam; PF00537; toxin_3; 1.
R    PRINTS; PR00284; TOXIN.
R    ProDom; PD000908; Scorpion_toxinL; 1.
R    SMART; SM00505; Knot1; 1.
W    Hypothetical protein.
Q    SEQUENCE   85 AA;  9426 MW;  6EFC463E02E95091 CRC64;

lignment Scores:
red. No.:              0.742          Length:         85
core:                  74.00          Matches:        12
ercent Similarity:     64.29%         Conservative:   6
est Local Similarity:  42.86%         Mismatches:     6
uery Match:            22.84%         Indels:         4
B:                     5              Gaps:           2

S-09-673-274B-2 (1-169) x Q9GUA7 (1-85)

y         52 TGCAACGGTGAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGC---------GGTTCC 102
             |||:::   |||||||||::::::|||   :::   |||:::|||         |||
b         41 CysAspAspGluCysLysLysLysGlyAlaGluSerGlyTyrCysGlnTrpAlaGlyVal 60

y        103 TTCGCTAACGTGAACTGCTGGTGC 126
             :::   |||      |||||||||
b         61 TyrGlyAsnAla---CysTrpCys 67

ESULT 15
9VZR3
D    Q9VZR3        PRELIMINARY;      PRT;   106 AA.
C    Q9VZR3;
T    01-MAY-2000 (TrEMBLrel. 13, Created)
T    01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
T    01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
E    CG10813 protein.
N    CG32279 OR CG10813.
S    Drosophila melanogaster (Fruit fly).
C    Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
C    Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
C    Ephydroidea; Drosophilidae; Drosophila.
X    NCBI_TaxID=7227;
N    [1]
P    SEQUENCE FROM N.A.
C    STRAIN=BERKELEY;
X    MEDLINE=20196006; PubMed=10731132;
A    Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
A    Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
A    George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
```

```
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003477; AAF47756.1; -.
DR   HSSP; P41964; 1MYN.
DR   FlyBase; FBgn0052279; CG32279.
DR   InterPro; IPR002118; Gamma-thionin.
DR   InterPro; IPR003614; Knot1.
DR   ProDom; PD002594; Gamma-thionin; 1.
DR   SMART; SM00505; Knot1; 1.
SQ   SEQUENCE   106 AA;  11598 MW;  A76479D6E44BC50C CRC64;

Alignment Scores:
Pred. No.:              0.875          Length:         106
Score:                  73.50          Matches:        16
Percent Similarity:     44.19%         Conservative:   3
Best Local Similarity:  37.21%         Mismatches:     17
Query Match:            22.69%         Indels:         7
DB:                     5              Gaps:           3

US-09-673-274B-2 (1-169) x Q9VZR3 (1-106)

Qy          4 AAGCTTATCGGTTCCTGC---GTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
              |||      |||   ||    :|:|||                        ||:
Db         68 LysTyrLysGlyProCysAlaValTrpAsp------------AsnGluMetCysArgArg 83

Qy         61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
                 :|||||      |||:::      .||||| ||           ::::::   |.|
Db         84 IleCysLysGluGluGlyHisIleSerGlyHisCysSer------ProSerLeuLysCys 101

Qy        121 TGGTGCGAG 129
              ||.:||||:|
Db        102 TrpCysGlu 104

Search completed: October 16, 2003, 17:16:12
Job time : 53.9701 secs
```

```
                 GenCore version 5.1.6
          Copyright (c) 1993 - 2003  Compugen Ltd.

M nucleic - protein search, using frame_plus_n2p model

un on:         October 16, 2003, 17:06:33 ; Search time 52.2159 Seconds
                                            (without alignments)
                                            1027.456 Million cell updates/sec

itle:          US-09-673-274B-2
erfect score:  324
equence:       1 gataagcttatcggttcctg..........gaacgtgtcgacggatccgg 169

coring table:  BLOSUM62
               Xgapop 10.0 , Xgapext  0.5
               Ygapop 10.0 , Ygapext  0.5
               Fgapop  6.0 , Fgapext  7.0
               Delop   6.0 , Delext   7.0

Searched:      1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters:      2215726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
 MODEL=frame+_n2p.model -DEV=xlp
 Q=/cgn2_1/USPTO_spool/US09673274/runat_16102003_170626_2523/app_query.fasta_1.654
-DB=A_Geneseq_19Jun03 -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
 LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
 MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
 USER=US09673274_@CGN_1_1_195_@runat_16102003_170626_2523 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
 DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
 FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :     A_Geneseq_19Jun03:*
               1:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
               2:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
               3:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
               4:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
               5:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
               6:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
               7:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
               8:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
               9:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
               10:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
               11:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
               12:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
               13:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
               14:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
               15:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
               16:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
               17:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
               18:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
               19:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
               20:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
               21:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
               22:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
               23:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
               24:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.
```

```
                                  SUMMARIES
                    %
Result            Query
  No.    Score    Match Length DB  ID                          Description
-------------------------------------------------------------------------------------
     1     267     82.4     44  21  AAY54375                    Amino acid sequenc
     2     267     82.4     44  23  ABB76860                    Antifungal/antibac
     3     267     82.4     44  23  ABB76868                    Antifungal/antibac
     4     267     82.4     49  21  AAY54374                    Amino acid sequenc
     5     267     82.4     74  21  AAY54376                    A fusion peptide o
     6     266     82.1     44  23  ABB76895                    Antifungal/antibac
     7     264     81.5     44  23  ABB76863                    Antifungal/antibac
     8     264     81.5     44  23  ABB76864                    Antifungal/antibac
     9     264     81.5     44  23  ABB76872                    Antifungal/antibac
    10     264     81.5     44  23  ABB76896                    Antifungal/antibac
    11     262     80.9     44  23  ABB76862                    Antifungal/antibac
    12     262     80.9     44  23  ABB76865                    Antifungal/antibac
    13     262     80.9     44  23  ABB76867                    Antifungal/antibac
    14     262     80.9     44  23  ABB76869                    Antifungal/antibac
    15     262     80.9     44  23  ABB76871                    Antifungal/antibac
    16     261     80.6     44  23  ABB76866                    Antifungal/antibac
    17     261     80.6     44  23  ABB76870                    Antifungal/antibac
    18     261     80.6     44  23  ABB76894                    Antifungal/antibac
    19     259     79.9     44  23  ABB76873                    Antifungal/antibac
    20     258     79.6     44  23  ABB76898                    Antifungal/antibac
    21     256     79.0     44  23  ABB76861                    Antifungal/antibac
    22     256     79.0     44  23  ABB76897                    Antifungal/antibac
    23     255     78.7     44  23  ABB76882                    Antifungal/antibac
    24     254     78.4     44  23  ABB76874                    Antifungal/antibac
    25     254     78.4     44  23  ABB76875                    Antifungal/antibac
    26     253     78.1     44  23  ABB76879                    Antifungal/antibac
    27     253     78.1     44  23  ABB76885                    Antifungal/antibac
    28     251     77.5     44  23  ABB76876                    Antifungal/antibac
    29     251     77.5     44  23  ABB76877                    Antifungal/antibac
    30     251     77.5     44  23  ABB76881                    Antifungal/antibac
    31     251     77.5     44  23  ABB76884                    Antifungal/antibac
    32     251     77.5     44  23  ABB76886                    Antifungal/antibac
    33     250     77.2     44  23  ABB76880                    Antifungal/antibac
    34     250     77.2     44  23  ABB76883                    Antifungal/antibac
    35     248     76.5     44  23  ABB76887                    Antifungal/antibac
    36     248     76.5     44  23  ABB76888                    Antifungal/antibac
    37     246     75.9     44  23  ABB76878                    Antifungal/antibac
    38     246     75.9     44  23  ABB76889                    Antifungal/antibac
    39     246     75.9     44  23  ABB76890                    Antifungal/antibac
    40     243     75.0     44  23  ABB76892                    Antifungal/antibac
    41     243     75.0     44  23  ABB76899                    Antifungal/antibac
    42     241     74.4     44  23  ABB76891                    Antifungal/antibac
    43     238     73.5     44  23  ABB76893                    Antifungal/antibac
    44     232     71.6     44  24  ABP55988                    Agrotis ipsilon Fu
    45     232     71.6     73  24  ABP55987                    Agrotis ipsilon Fu


                                  ALIGNMENTS

RESULT 1
AAY54375
ID   AAY54375 standard; Peptide; 44 AA.
XX
AC   AAY54375;
XX
DT   06-APR-2000  (first entry)
XX
DE   Amino acid sequence of a heliomycin peptide.
XX
KW   Heliomycin; polycysteine peptide; antifungal; antibacterial;
KW   fungal infection; plant protection; animal protection;
KW   transgenic plant; fungi; bacteria; Cercospora beticola;
KW   Cladosporium herbarum; Fusarium culmorum; F. oxysporum;
KW   Phytophthora cinnamoni.
XX
OS   Heliothis virescens.
XX
```

```
PN   FR2777568-A1.
XX
PD   22-OCT-1999.
XX
PF   15-APR-1998;    98FR-0004933.
XX
PR   15-APR-1998;    98FR-0004933.
XX
PA   (RHON ) RHONE-POULENC AGROCHIMIE.
XX
DR   WPI; 2000-108532/10.
DR   N-PSDB; AAZ45692.
XX
PT   New polycysteine peptides, designated heliomycin, with antifungal or
PT   antibacterial activity, useful in medicine or plant protection  -
XX
PS   Claim 7; Page 31-32; 46pp; French.
XX
CC   The present sequence represents the heliomycin peptide. Heliomycin is a
CC   polycysteine peptide with antifungal and antibacterial activity,
CC   isolated from the lepidopteron Heliothis virescens. Heliomycin peptides
CC   contain the sequence given in AAY54378. The heliomycin peptides are
CC   used as pharmaceuticals for treating or preventing fungal infections
CC   in humans and animals. They may also be used as antifungal agents for
CC   plant protection. Nucleic acid encoding the heliomycin peptides is
CC   used to generate transgenic plants that are resistant to some fungi
CC   and bacteria (specifically Cercospora beticola, Cladosporium herbarum,
CC   Fusarium culmorum, F. oxysporum and Phytophthora cinnamoni), and for
CC   recombinant production of the peptides.
XX
SQ   Sequence   44 AA;

Alignment Scores:
Pred. No.:              6.42e-26       Length:        44
Score:                  267.00         Matches:       44
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            82.41%         Indels:        0
DB:                     21             Gaps:          0

US-09-673-274B-2 (1-169) x AAY54375 (1-44)

Qy          1 GATAAGCTTATCGGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AspLysLeuIleGlySerCysValTrpGlyAlaValAsnTyrThrSerAspCysAsnGly 20

Qy         61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 GluCysLysArgArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnValAsnCys 40

Qy        121 TGGTGCGAGACT 132
              ||||||||||||
Db         41 TrpCysGluThr 44

RESULT 2
ABB76860
ID   ABB76860 standard; Protein; 44 AA.
XX
AC   ABB76860;
XX
DT   03-JUL-2002  (first entry)
XX
DE   Antifungal/antibacterial peptide Helio.
XX
KW   Antifungal; antibacterial; fungicide; heliomicine; human medicine;
KW   veterinary medicine; plant protection; bacterial infection;
KW   fungal infection.
XX
OS   Heliothis virescens.
XX
PN   WO200206324-A2.
XX
PD   24-JAN-2002.
XX
PF   05-JUL-2001; 2001WO-FR02164.
XX
PR   13-JUL-2000; 2000FR-0009248.
PR   19-SEP-2000; 2000FR-0011949.
XX
PA   (ENTO-) ENTOMED SA.
XX
PI   Dimarcq J,  Legrain M,  Menin L;
XX
DR   WPI; 2002-179779/23.
XX
PT   New peptide derivatives of heliomicine, useful as antibacterial and
PT   antifungal agents in human or veterinary medicine and agriculture  -
XX
PS   Claim 15; Page 57; 100pp; French.
XX
CC   The present invention relates to peptide derivatives of heliomicine with
CC   antifungal and/or antibacterial activity. The present sequence is one
CC   such peptide. The peptides are useful in human and veterinary medicine
CC   and in plant protection, for control of bacterial and fungal infections.
CC   An antifungal and/or antibacterial composition containing at least one
CC   peptide is also claimed.
XX
SQ   Sequence   44 AA;

Alignment Scores:
Pred. No.:              6.42e-26       Length:        44
Score:                  267.00         Matches:       44
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            82.41%         Indels:        0
DB:                     23             Gaps:          0

US-09-673-274B-2 (1-169) x ABB76860 (1-44)

Qy          1 GATAAGCTTATCGGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AspLysLeuIleGlySerCysValTrpGlyAlaValAsnTyrThrSerAspCysAsnGly 20

Qy         61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 GluCysLysArgArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnValAsnCys 40

Qy        121 TGGTGCGAGACT 132
              ||||||||||||
Db         41 TrpCysGluThr 44

RESULT 3
ABB76868
ID   ABB76868 standard; Protein; 44 AA.
XX
AC   ABB76868;
XX
DT   03-JUL-2002  (first entry)
XX
DE   Antifungal/antibacterial peptide pEM23.
XX
KW   Antifungal; antibacterial; fungicide; heliomicine; human medicine;
KW   veterinary medicine; plant protection; bacterial infection;
KW   fungal infection.
XX
OS   Synthetic.
XX
PN   WO200206324-A2.
XX
PD   24-JAN-2002.
XX
PF   05-JUL-2001; 2001WO-FR02164.
XX
PR   13-JUL-2000; 2000FR-0009248.
PR   19-SEP-2000; 2000FR-0011949.
XX
```

```
A   (ENTO-) ENTOMED SA.
X
I   Dimarcq J,  Legrain M,  Menin L;
X
R   WPI; 2002-179779/23.
X
T   New peptide derivatives of heliomicine, useful as antibacterial and
T   antifungal agents in human or veterinary medicine and agriculture  -
X
S   Claim 15; Page 57; 100pp; French.
X
C   The present invention relates to peptide derivatives of heliomicine with
C   antifungal and/or antibacterial activity. The present sequence is one
C   such peptide. The peptides are useful in human and veterinary medicine
C   and in plant protection, for control of bacterial and fungal infections.
C   An antifungal and/or antibacterial composition containing at least one
C   peptide is also claimed.
X
Q   Sequence   44 AA;

lignment Scores:
red. No.:                6.42e-26      Length:         44
core:                    267.00        Matches:        44
ercent Similarity:       100.00%       Conservative:   0
est Local Similarity:    100.00%       Mismatches:     0
uery Match:              82.41%        Indels:         0
B:                       23            Gaps:           0

S-09-673-274B-2 (1-169) x ABB76868 (1-44)

y          1 GATAAGCTTATCGGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b          1 AspLysLeuIleGlySerCysValTrpGlyAlaValAsnTyrThrSerAspCysAsnGly 20

y         61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         21 GluCysLysArgArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnValAsnCys 40

y        121 TGGTGCGAGACT 132
               ||||||||||||
b         41 TrpCysGluThr 44

ESULT 4
AY54374
D   AAY54374 standard; Peptide; 49 AA.
X
C   AAY54374;
X
T   06-APR-2000  (first entry)
X
E   Amino acid sequence of a fusion peptide of MF-alpha1/heliomycin.
X
W   Mat alpha1 factor; MF-alpha1; heliomycin; polycysteine peptide;
W   antifungal; antibacterial; fungal infection; plant protection;
W   animal protection; transgenic plant; fungi; bacteria;
W   Cercospora beticola; Cladosporium herbarum; Fusarium culmorum;
W   F. oxysporum; Phytophthora cinnamoni.
X
S   Chimeric - Saccharomyces cerevisiae.
S   Chimeric - Heliothis virescens.
X
H   Key             Location/Qualifiers
T   Peptide         1..5
T                   /note= "residues derived from the yeast Mat-alpha1
T                          (MF-alpha1) factor"
T   Peptide         6..49
T                   /note= "heliomycin residues"
X
N   FR2777568-A1.
X
D   22-OCT-1999.
X
F   15-APR-1998;   98FR-0004933.
```

```
XX
PR   15-APR-1998;   98FR-0004933.
XX
PA   (RHON ) RHONE-POULENC AGROCHIMIE.
XX
DR   WPI; 2000-108532/10.
DR   N-PSDB; AAZ45691.
XX
PT   New polycysteine peptides, designated heliomycin, with antifungal or
PT   antibacterial activity, useful in medicine or plant protection  -
XX
PS   Claim 13; Page 31; 46pp; French.
XX
CC   The present sequence represents a fusion peptide of 5 amino acids
CC   derived from the Mat alpha1 factor (MF-alpha1) and heliomycin. The
CC   attachment of the MF-alpha1 sequence to the N-terminal aids secretion
CC   of the heliomycin peptide. Heliomycin is a polycysteine peptide with
CC   antifungal and antibacterial activity, isolated from the lepidopteron
CC   Heliothis virescens. Heliomycin peptides contain the sequence given in
CC   AAY54378. The heliomycin peptides are used as pharmaceuticals for
CC   treating or preventing fungal infections in humans and animals. They
CC   may also be used as antifungal agents for plant protection. Nucleic
CC   acid encoding the heliomycin peptides is used to generate transgenic
CC   plants that are resistant to some fungi and bacteria (specifically
CC   Cercospora beticola, Cladosporium herbarum, Fusarium culmorum,
CC   F. oxysporum and Phytophthora cinnamoni), and for recombinant production
CC   of the peptides.
XX
SQ   Sequence   49 AA;

Alignment Scores:
Pred. No.:                6.54e-26      Length:         49
Score:                    267.00        Matches:        44
Percent Similarity:       100.00%       Conservative:   0
Best Local Similarity:    100.00%       Mismatches:     0
Query Match:              82.41%        Indels:         0
DB:                       21            Gaps:           0

US-09-673-274B-2 (1-169) x AAY54374 (1-49)

Qy          1 GATAAGCTTATCGGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6 AspLysLeuIleGlySerCysValTrpGlyAlaValAsnTyrThrSerAspCysAsnGly 25

Qy         61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 GluCysLysArgArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnValAsnCys 45

Qy        121 TGGTGCGAGACT 132
              ||||||||||||
Db         46 TrpCysGluThr 49

RESULT 5
AAY54376
ID   AAY54376 standard; Peptide; 74 AA.
XX
AC   AAY54376;
XX
DT   06-APR-2000  (first entry)
XX
DE   A fusion peptide of PR-1alpha signal peptide/heliomycin.
XX
KW   PR-1alpha; heliomycin; polycysteine peptide; antifungal; antibacterial;
KW   fungal infection; plant protection; animal protection; transgenic plant;
KW   fungi; bacteria; Cercospora beticola; Cladosporium herbarum;
KW   Fusarium culmorum; F. oxysporum; Phytophthora cinnamoni.
XX
OS   Chimeric - Nicotiana sp.
OS   Chimeric - Heliothis virescens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..30
FT                   /note= "tobacco PR-1alpha signal peptide"
```

```
FT   Peptide         31..74
FT                   /note= "heliomycin"
XX
PN   FR2777568-A1.
XX
PD   22-OCT-1999.
XX
PF   15-APR-1998;   98FR-0004933.
XX
PR   15-APR-1998;   98FR-0004933.
XX
PA   (RHON ) RHONE-POULENC AGROCHIMIE.
XX
DR   WPI; 2000-108532/10.
DR   N-PSDB; AAZ45693.
XX
PT   New polycysteine peptides, designated heliomycin, with antifungal or
PT   antibacterial activity, useful in medicine or plant protection  -
XX
PS   Claim 13; Page 32; 46pp; French.
XX
CC   The present sequence represents a fusion peptide of the tobacco
CC   PR-1alpha signal peptide and heliomycin. The attachment of the
CC   PR-1alpha sequence to the N-terminal aids secretion of the heliomycin
CC   peptide. Heliomycin is a polycysteine peptide with antifungal and
CC   antibacterial activity, isolated from the lepidopteron Heliothis
CC   virescens. Heliomycin peptides contain the sequence given in
CC   AAY54378. The heliomycin peptides are used as pharmaceuticals for
CC   treating or preventing fungal infections in humans and animals. They
CC   may also be used as antifungal agents for plant protection. Nucleic
CC   acid encoding the heliomycin peptides is used to generate transgenic
CC   plants that are resistant to some fungi and bacteria (specifically
CC   Cercospora beticola, Cladosporium herbarum, Fusarium culmorum,
CC   F. oxysporum and Phytophthora cinnamonii, and for recombinant production
CC   of the peptides.
XX
SQ   Sequence   74 AA;

Alignment Scores:
Pred. No.:              7.01e-26      Length:        74
Score:                  267.00        Matches:       44
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            82.41%        Indels:        0
DB:                     21            Gaps:          0

US-09-673-274B-2 (1-169) x AAY54376 (1-74)

Qy          1 GATAAGCTTATCGGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         31 AspLysLeuIleGlySerCysValTrpGlyAlaValAsnTyrThrSerAspCysAsnGly 50

Qy         61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         51 GluCysLysArgArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnValAsnCys 70

Qy        121 TGGTGCGAGACT 132
              ||||||||||||
Db         71 TrpCysGluThr 74

RESULT 6
ABB76895
ID   ABB76895 standard; Protein; 44 AA.
XX
AC   ABB76895;
XX
DT   03-JUL-2002  (first entry)
XX
DE   Antifungal/antibacterial peptide #2.
XX
KW   Antifungal; antibacterial; fungicide; heliomicine; human medicine;
KW   veterinary medicine; plant protection; bacterial infection;
KW   fungal infection.
XX
OS   Synthetic.
XX
PN   WO200206324-A2.
XX
PD   24-JAN-2002.
XX
PF   05-JUL-2001; 2001WO-FR02164.
XX
PR   13-JUL-2000; 2000FR-0009248.
PR   19-SEP-2000; 2000FR-0011949.
XX
PA   (ENTO-) ENTOMED SA.
XX
PI   Dimarcq J,  Legrain M,  Menin L;
XX
DR   WPI; 2002-179779/23.
XX
PT   New peptide derivatives of heliomicine, useful as antibacterial and
PT   antifungal agents in human or veterinary medicine and agriculture  -
XX
PS   Disclosure; Page 86; 100pp; French.
XX
CC   The present invention relates to peptide derivatives of heliomicine with
CC   antifungal and/or antibacterial activity. The present sequence is one
CC   such peptide. The peptides are useful in human and veterinary medicine
CC   and in plant protection, for control of bacterial and fungal infections.
CC   An antifungal and/or antibacterial composition containing at least one
CC   peptide is also claimed.
XX
SQ   Sequence   44 AA;

Alignment Scores:
Pred. No.:              8.65e-26      Length:        44
Score:                  266.00        Matches:       43
Percent Similarity:     100.00%       Conservative:  1
Best Local Similarity:  97.73%        Mismatches:    0
Query Match:            82.10%        Indels:        0
DB:                     23            Gaps:          0

US-09-673-274B-2 (1-169) x ABB76895 (1-44)

Qy          1 GATAAGCTTATCGGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AspLysLeuIleGlySerCysValTrpGlyAlaValAsnTyrThrSerAspCysAsnGly 20

Qy         61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||:::||||||
Db         21 GluCysLysArgArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnIleAsnCys 40

Qy        121 TGGTGCGAGACT 132
              ||||||||||||
Db         41 TrpCysGluThr 44

RESULT 7
ABB76863
ID   ABB76863 standard; Protein; 44 AA.
XX
AC   ABB76863;
XX
DT   03-JUL-2002  (first entry)
XX
DE   Antifungal/antibacterial peptide pEM38.
XX
KW   Antifungal; antibacterial; fungicide; heliomicine; human medicine;
KW   veterinary medicine; plant protection; bacterial infection;
KW   fungal infection.
XX
OS   Synthetic.
XX
PN   WO200206324-A2.
XX
PD   24-JAN-2002.
```

```
X
F    05-JUL-2001; 2001WO-FR02164.
X
R    13-JUL-2000; 2000FR-0009248.
R    19-SEP-2000; 2000FR-0011949.
X
A    (ENTO-) ENTOMED SA.
X
I    Dimarcq J,  Legrain M,  Menin L;
X
R    WPI; 2002-179779/23.
X
T    New peptide derivatives of heliomicine, useful as antibacterial and
T    antifungal agents in human or veterinary medicine and agriculture  -
X
S    Claim 15; Page 57; 100pp; French.
X
C    The present invention relates to peptide derivatives of heliomicine with
C    antifungal and/or antibacterial activity. The present sequence is one
C    such peptide. The peptides are useful in human and veterinary medicine
C    and in plant protection, for control of bacterial and fungal infections.
C    An antifungal and/or antibacterial composition containing at least one
C    peptide is also claimed.
X
Q    Sequence   44 AA;

lignment Scores:
red. No.:                1.57e-25        Length:         44
core:                    264.00          Matches:        43
ercent Similarity:       100.00%         Conservative:   1
est Local Similarity:    97.73%          Mismatches:     0
uery Match:              81.48%          Indels:         0
B:                       23              Gaps:           0

S-09-673-274B-2 (1-169) x ABB76863 (1-44)

y          1 GATAAGCTTATCGGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
               |||||||||||||||:::||||||||||||||||||||||||||||||||||||||||||
b          1 AspLysLeuIleGlyThrCysValTrpGlyAlaValAsnTyrThrSerAspCysAsnGly 20

y         61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
               ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||
b         21 GluCysLysArgArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnValAsnCys 40

y        121 TGGTGCGAGACT 132
               ||||||||||||
b         41 TrpCysGluThr 44

ESULT 8
BB76864
D    ABB76864 standard; Protein; 44 AA.
X
C    ABB76864;
X
T    03-JUL-2002  (first entry)
X
E    Antifungal/antibacterial peptide pEM43.
X
W    Antifungal; antibacterial; fungicide; heliomicine; human medicine;
W    veterinary medicine; plant protection; bacterial infection;
W    fungal infection.
X
S    Synthetic.
X
N    WO200206324-A2.
X
D    24-JAN-2002.
X
F    05-JUL-2001; 2001WO-FR02164.
X
R    13-JUL-2000; 2000FR-0009248.
R    19-SEP-2000; 2000FR-0011949.
X
```

```
PA   (ENTO-) ENTOMED SA.
XX
PI   Dimarcq J,  Legrain M,  Menin L;
XX
DR   WPI; 2002-179779/23.
XX
PT   New peptide derivatives of heliomicine, useful as antibacterial and
PT   antifungal agents in human or veterinary medicine and agriculture  -
XX
PS   Claim 15; Page 57; 100pp; French.
XX
CC   The present invention relates to peptide derivatives of heliomicine with
CC   antifungal and/or antibacterial activity. The present sequence is one
CC   such peptide. The peptides are useful in human and veterinary medicine
CC   and in plant protection, for control of bacterial and fungal infections.
CC   An antifungal and/or antibacterial composition containing at least one
CC   peptide is also claimed.
XX
SQ   Sequence   44 AA;

Alignment Scores:
Pred. No.:                1.57e-25        Length:         44
Score:                    264.00          Matches:        43
Percent Similarity:       100.00%         Conservative:   1
Best Local Similarity:    97.73%          Mismatches:     0
Query Match:              81.48%          Indels:         0
DB:                       23              Gaps:           0

US-09-673-274B-2 (1-169) x ABB76864 (1-44)

Qy          1 GATAAGCTTATCGGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
                |:||||||:|||||:|||| :!||| |'||||||!||||||:|||||||:::||||:! ||||
Db          1 AspLysLeuIleGlySerCysValTrpGlyAlaValAsnTyrThrThrAspCysAsnGly 20

Qy         61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
                |:|||||||:||||:|!|||| :||||:'||||||:|·||||:|:||||||'|!|||:'|||| ·:
Db         21 GluCysLysArgArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnValAsnCys 40

Qy        121 TGGTGCGAGACT 132
                |||:||||:' |||
Db         41 TrpCysGluThr 44

RESULT 9
ABB76872
ID   ABB76872 standard; Protein; 44 AA.
XX
AC   ABB76872;
XX
DT   03-JUL-2002  (first entry)
XX
DE   Antifungal/antibacterial peptide pEM21.
XX
KW   Antifungal; antibacterial; fungicide; heliomicine; human medicine;
KW   veterinary medicine; plant protection; bacterial infection;
KW   fungal infection.
XX
OS   Synthetic.
XX
PN   WO200206324-A2.
XX
PD   24-JAN-2002.
XX
PF   05-JUL-2001; 2001WO-FR02164.
XX
PR   13-JUL-2000; 2000FR-0009248.
PR   19-SEP-2000; 2000FR-0011949.
XX
PA   (ENTO-) ENTOMED SA.
XX
PI   Dimarcq J,  Legrain M,  Menin L;
XX
DR   WPI; 2002-179779/23.
XX
```

```
PT   New peptide derivatives of heliomicine, useful as antibacterial and
PT   antifungal agents in human or veterinary medicine and agriculture  -
XX
PS   Claim 15; Page 57; 100pp; French.
XX
CC   The present invention relates to peptide derivatives of heliomicine with
CC   antifungal and/or antibacterial activity. The present sequence is one
CC   such peptide. The peptides are useful in human and veterinary medicine
CC   and in plant protection, for control of bacterial and fungal infections.
CC   An antifungal and/or antibacterial composition containing at least one
CC   peptide is also claimed.
XX
SQ   Sequence   44 AA;

Alignment Scores:
Pred. No.:                1.57e-25       Length:          44
Score:                    264.00         Matches:         43
Percent Similarity:       100.00%        Conservative:    1
Best Local Similarity:    97.73%         Mismatches:      0
Query Match:              81.48%         Indels:          0
DB:                       23             Gaps:            0

US-09-673-274B-2 (1-169) x ABB76872 (1-44)

Qy          1 GATAAGCTTATCGGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AspLysLeuIleGlySerCysValTrpGlyAlaValAsnTyrThrSerAspCysAsnGly 20

Qy         61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 GluCysLysArgArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnValAsnCys 40

Qy        121 TGGTGCGAGACT 132
              ||||||:::|||
Db         41 TrpCysGlnThr 44

RESULT 10
ABB76896
ID   ABB76896 standard; Protein; 44 AA.
XX
AC   ABB76896;
XX
DT   03-JUL-2002  (first entry)
XX
DE   Antifungal/antibacterial peptide #3.
XX
KW   Antifungal; antibacterial; fungicide; heliomicine; human medicine;
KW   veterinary medicine; plant protection; bacterial infection;
KW   fungal infection.
XX
OS   Synthetic.
XX
PN   WO200206324-A2.
XX
PD   24-JAN-2002.
XX
PF   05-JUL-2001; 2001WO-FR02164.
XX
PR   13-JUL-2000; 2000FR-0009248.
PR   19-SEP-2000; 2000FR-0011949.
XX
PA   (ENTO-) ENTOMED SA.
XX
PI   Dimarcq J,  Legrain M,  Menin L;
XX
DR   WPI; 2002-179779/23.
XX
PT   New peptide derivatives of heliomicine, useful as antibacterial and
PT   antifungal agents in human or veterinary medicine and agriculture  -
XX
PS   Disclosure; Page 88; 100pp; French.
XX
CC   The present invention relates to peptide derivatives of heliomicine with
CC   antifungal and/or antibacterial activity. The present sequence is one
CC   such peptide. The peptides are useful in human and veterinary medicine
CC   and in plant protection, for control of bacterial and fungal infections.
CC   An antifungal and/or antibacterial composition containing at least one
CC   peptide is also claimed.
XX
SQ   Sequence   44 AA;

Alignment Scores:
Pred. No.:                1.57e-25       Length:          44
Score:                    264.00         Matches:         43
Percent Similarity:       100.00%        Conservative:    1
Best Local Similarity:    97.73%         Mismatches:      0
Query Match:              81.48%         Indels:          0
DB:                       23             Gaps:            0

US-09-673-274B-2 (1-169) x ABB76896 (1-44)

Qy          1 GATAAGCTTATCGGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AspLysLeuIleGlySerCysValTrpGlyAlaValAsnTyrThrSerAspCysAsnGly 20

Qy         61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 GluCysLysArgArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnValAsnCys 40

Qy        121 TGGTGCGAGACT 132
              ||||||:::|||
Db         41 TrpCysGlnThr 44

RESULT 11
ABB76862
ID   ABB76862 standard; Protein; 44 AA.
XX
AC   ABB76862;
XX
DT   03-JUL-2002  (first entry)
XX
DE   Antifungal/antibacterial peptide pEM37.
XX
KW   Antifungal; antibacterial; fungicide; heliomicine; human medicine;
KW   veterinary medicine; plant protection; bacterial infection;
KW   fungal infection.
XX
OS   Synthetic.
XX
PN   WO200206324-A2.
XX
PD   24-JAN-2002.
XX
PF   05-JUL-2001; 2001WO-FR02164.
XX
PR   13-JUL-2000; 2000FR-0009248.
PR   19-SEP-2000; 2000FR-0011949.
XX
PA   (ENTO-) ENTOMED SA.
XX
PI   Dimarcq J,  Legrain M,  Menin L;
XX
DR   WPI; 2002-179779/23.
XX
PT   New peptide derivatives of heliomicine, useful as antibacterial and
PT   antifungal agents in human or veterinary medicine and agriculture  -
XX
PS   Claim 15; Page 57; 100pp; French.
XX
CC   The present invention relates to peptide derivatives of heliomicine with
CC   antifungal and/or antibacterial activity. The present sequence is one
CC   such peptide. The peptides are useful in human and veterinary medicine
CC   and in plant protection, for control of bacterial and fungal infections.
CC   An antifungal and/or antibacterial composition containing at least one
CC   peptide is also claimed.
XX
```

```
Q   Sequence   44 AA;

lignment Scores:
red. No.:                2.84e-25       Length:        44
core:                    262.00         Matches:       43
ercent Similarity:       100.00%        Conservative:  1
est Local Similarity:    97.73%         Mismatches:    0
uery Match:              80.86%         Indels:        0
B:                       23             Gaps:          0

S-09-673-274B-2 (1-169) x ABB76862 (1-44)

y          1 GATAAGCTTATCGGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
               :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b          1 AsnLysLeuIleGlySerCysValTrpGlyAlaValAsnTyrThrSerAspCysAsnGly 20

y         61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         21 GluCysLysArgArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnValAsnCys 40

y        121 TGGTGCGAGACT 132
             ||||||||||||
b         41 TrpCysGluThr 44

ESULT 12
BB76865
D   ABB76865 standard; Protein; 44 AA.
X
C   ABB76865;
X
T   03-JUL-2002  (first entry)
X
E   Antifungal/antibacterial peptide pEM42.
X
W   Antifungal; antibacterial; fungicide; heliomicine; human medicine;
W   veterinary medicine; plant protection; bacterial infection;
W   fungal infection.
X
S   Synthetic.
X
N   WO200206324-A2.
X
D   24-JAN-2002.
X
F   05-JUL-2001; 2001WO-FR02164.
X
R   13-JUL-2000; 2000FR-0009248.
R   19-SEP-2000; 2000FR-0011949.
X
A   (ENTO-) ENTOMED SA.
X
I   Dimarcq J,  Legrain M,  Menin L;
X
R   WPI; 2002-179779/23.
X
T   New peptide derivatives of heliomicine, useful as antibacterial and
T   antifungal agents in human or veterinary medicine and agriculture  -
X
PS  Claim 15; Page 57; 100pp; French.
X
CC  The present invention relates to peptide derivatives of heliomicine with
CC  antifungal and/or antibacterial activity. The present sequence is one
CC  such peptide. The peptides are useful in human and veterinary medicine
CC  and in plant protection, for control of bacterial and fungal infections.
CC  An antifungal and/or antibacterial composition containing at least one
CC  peptide is also claimed.
XX
SQ  Sequence   44 AA;

Alignment Scores:
Pred. No.:               2.84e-25       Length:        44
Score:                   262.00         Matches:       43
Percent Similarity:      97.73%         Conservative:  0
```

```
Best Local Similarity:   97.73%         Mismatches:    1
Query Match:             80.86%         Indels:        0
DB:                      23             Gaps:          0

US-09-673-274B-2 (1-169) x ABB76865 (1-44)

Qy          1 GATAAGCTTATCGGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
              |||||||||||||||||||||||||||||||||||||||||||||   |||||||||||:
Db          1 AspLysLeuIleGlySerCysValTrpGlyAlaValAsnTyrThrArgAspCysAsnGly 20

Qy         61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db         21 GluCysLysArgArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnValAsnCys 40

Qy        121 TGGTGCGAGACT 132
              ||||||||||||
Db         41 TrpCysGluThr 44

RESULT 13
ABB76867
ID   ABB76867 standard; Protein; 44 AA.
XX
AC   ABB76867;
XX
DT   03-JUL-2002  (first entry)
XX
DE   Antifungal/antibacterial peptide pEM22.
XX
KW   Antifungal; antibacterial; fungicide; heliomicine; human medicine;
KW   veterinary medicine; plant protection; bacterial infection;
KW   fungal infection.
XX
OS   Synthetic.
XX
PN   WO200206324-A2.
XX
PD   24-JAN-2002.
XX
PF   05-JUL-2001; 2001WO-FR02164.
XX
PR   13-JUL-2000; 2000FR-0009248.
PR   19-SEP-2000; 2000FR-0011949.
XX
PA   (ENTO-) ENTOMED SA.
XX
PI   Dimarcq J,  Legrain M,  Menin L;
XX
DR   WPI; 2002-179779/23.
XX
PT   New peptide derivatives of heliomicine, useful as antibacterial and
PT   antifungal agents in human or veterinary medicine and agriculture
XX
PS   Claim 15; Page 57; 100pp; French.
XX
CC   The present invention relates to peptide derivatives of heliomicine with
CC   antifungal and/or antibacterial activity. The present sequence is one
CC   such peptide. The peptides are useful in human and veterinary medicine
CC   and in plant protection, for control of bacterial and fungal infections.
CC   An antifungal and/or antibacterial composition containing at least one
CC   peptide is also claimed.
XX
SQ   Sequence   44 AA;

Alignment Scores:
Pred. No.:               2.84e-25       Length:        44
Score:                   262.00         Matches:       43
Percent Similarity:      97.73%         Conservative:  0
Best Local Similarity:   97.73%         Mismatches:    1
Query Match:             80.86%         Indels:        0
DB:                      23             Gaps:          0

US-09-673-274B-2 (1-169) x ABB76867 (1-44)
```

```
Qy          1 GATAAGCTTATCGGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
              |:||||||||||||||:||||||||:|||||:|||||:|||||:'||||:||||||||:||
Db          1 AspLysLeuIleGlySerCysValTrpGlyAlaValAsnTyrThrSerAspCysAsnGly 20

Qy         61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
              |:|:|||||||:|'|||||||:||||||:||||:||||:|:||||'|||    |||:|||:|||
Db         21 GluCysLysArgArgGlyTyrLysGlyGlyHisCysGlySerPheIleAsnValAsnCys 40

Qy        121 TGGTGCGAGACT 132
              ·|'|||||:|'|
Db         41 TrpCysGluThr 44

RESULT 14
ABB76869
ID   ABB76869 standard; Protein; 44 AA.
XX
AC   ABB76869;
XX
DT   03-JUL-2002  (first entry)
XX
DE   Antifungal/antibacterial peptide pEM25.
XX
KW   Antifungal; antibacterial; fungicide; heliomicine; human medicine;
KW   veterinary medicine; plant protection; bacterial infection;
KW   fungal infection.
XX
OS   Synthetic.
XX
PN   WO200206324-A2.
XX
PD   24-JAN-2002.
XX
PF   05-JUL-2001; 2001WO-FR02164.
XX
PR   13-JUL-2000; 2000FR-0009248.
PR   19-SEP-2000; 2000FR-0011949.
XX
PA   (ENTO-) ENTOMED SA.
XX
PI   Dimarcq J,  Legrain M,  Menin L;
XX
DR   WPI; 2002-179779/23.
XX
PT   New peptide derivatives of heliomicine, useful as antibacterial and
PT   antifungal agents in human or veterinary medicine and agriculture  -
XX
PS   Claim 15; Page 57; 100pp; French.
XX
CC   The present invention relates to peptide derivatives of heliomicine with
CC   antifungal and/or antibacterial activity. The present sequence is one
CC   such peptide. The peptides are useful in human and veterinary medicine
CC   and in plant protection, for control of bacterial and fungal infections.
CC   An antifungal and/or antibacterial composition containing at least one
CC   peptide is also claimed.
XX
SQ   Sequence   44 AA;

Alignment Scores:
Pred. No.:              2.84e-25      Length:         44
Score:                  262.00        Matches:        43
Percent Similarity:     97.73%        Conservative:   0
Best Local Similarity:  97.73%        Mismatches:     1
Query Match:            80.86%        Indels:         0
DB:                     23            Gaps:           0

US-09-673-274B-2 (1-169) x ABB76869 (1-44)

Qy          1 GATAAGCTTATCGGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
              :|'||||||||||||||:||||||||||:|||||:|||||'||||||||:||||||||||||
Db          1 AspLysLeuIleGlySerCysValTrpGlyAlaValAsnTyrThrSerAspCysAsnGly 20

Qy         61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
              |:|||||||:|||||||||||||:|||||.|||:||||:.||||'    ':||||:||||:
```

```
Db         21 GluCysLysArgArgGlyTyrLysGlyGlyHisCysGlySerPheLeuAsnValAsnCys 40

Qy        121 TGGTGCGAGACT 132
              ||||||||:||
Db         41 TrpCysGluThr 44

RESULT 15
ABB76871
ID   ABB76871 standard; Protein; 44 AA.
XX
AC   ABB76871;
XX
DT   03-JUL-2002  (first entry)
XX
DE   Antifungal/antibacterial peptide pEM7.
XX
KW   Antifungal; antibacterial; fungicide; heliomicine; human medicine;
KW   veterinary medicine; plant protection; bacterial infection;
KW   fungal infection.
XX
OS   Synthetic.
XX
PN   WO200206324-A2.
XX
PD   24-JAN-2002.
XX
PF   05-JUL-2001; 2001WO-FR02164.
XX
PR   13-JUL-2000; 2000FR-0009248.
PR   19-SEP-2000; 2000FR-0011949.
XX
PA   (ENTO-) ENTOMED SA.
XX
PI   Dimarcq J,  Legrain M,  Menin L;
XX
DR   WPI; 2002-179779/23.
XX
PT   New peptide derivatives of heliomicine, useful as antibacterial and
PT   antifungal agents in human or veterinary medicine and agriculture  -
XX
PS   Claim 15; Page 57; 100pp; French.
XX
CC   The present invention relates to peptide derivatives of heliomicine with
CC   antifungal and/or antibacterial activity. The present sequence is one
CC   such peptide. The peptides are useful in human and veterinary medicine
CC   and in plant protection, for control of bacterial and fungal infections.
CC   An antifungal and/or antibacterial composition containing at least one
CC   peptide is also claimed.
XX
SQ   Sequence   44 AA;

Alignment Scores:
Pred. No.:              2.84e-25      Length:         44
Score:                  262.00        Matches:        43
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            80.86%        Indels:         0
DB:                     23.           Gaps:           0

US-09-673-274B-2 (1-169) x ABB76871 (1-44)

Qy          1 GATAAGCTTATCGGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AspLysLeuIleGlySerCysValTrpGlyAlaValAsnTyrThrSerAspCysAsnGly 20

Qy         61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
              |||||:||||:|||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 GluCysLysArgArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnValAsnCys 40

Qy        121 TGGTGCGAG 129
              ||:||||:|
Db         41 TrpCysGlu 43
```

```
earch completed: October 16, 2003, 17:10:35
ob time : 52.2159 secs
```

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2003  Compugen Ltd.

M nucleic - protein search, using frame_plus_n2p model

un on:          October 16, 2003, 17:12:07 ; Search time 35.0914 Seconds
                                         (without alignments)
                                         1578.850 Million cell updates/sec

itle:           US-09-673-274B-2
erfect score:   324
equence:        1 gataagcttatcggttcctg..........gaacgtgtcgacggatccgg 169

coring table:   BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

earched:        609560 seqs, 163917102 residues

otal number of hits satisfying chosen parameters:      1219120

inimum DB seq length: 0
aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%
                Maximum Match 100%
                Listing first 45 summaries

ommand line parameters:
 MODEL=frame+_n2p.model -DEV=xlp
 Q=/cgn2_1/USPTO_spool/US09673274/runat_16102003_170630_2743/app_query.fasta_1.654
 DB=Published_Applications_AA -QFMT=fastan -SUFFIX=rapb -MINMATCH=0.1
 LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
 TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
 THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
 MAXLEN=2000000000 -USER=US09673274_@CGN_1_1_106_@runat_16102003_170630_2743
 NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
 LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
 FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

atabase :       Published_Applications_AA:*
                1:   /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
                2:   /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
                3:   /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
                4:   /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
                5:   /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
                6:   /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
                7:   /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
                8:   /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
                9:   /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
                10:   /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
                11:   /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
                12:   /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
                13:   /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
                14:   /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
                15:   /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
                16:   /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
                17:   /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
                18:   /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                  SUMMARIES
                %
esult           Query
 No.    Score   Match Length DB  ID                            Description
------------------------------------------------------------------------------
      1     232    71.6     44  15  US-10-125-258-119          Sequence 119, App
      2     232    71.6     73  15  US-10-125-258-117          Sequence 117, App
      3    79.5    24.5     65  15  US-10-264-480-13           Sequence 13, Appl
      4    79.5    24.5     66  15  US-10-264-480-11           Sequence 11, Appl
      5    76.5    23.6     61  15  US-10-264-480-9            Sequence 9, Appli
      6      75    23.1     61  9  US-09-829-481-2            Sequence 2, Appli
      7    74.5    23.0     44  12  US-10-180-247-4            Sequence 4, Appli
      8    74.5    23.0     69  12  US-10-180-247-6            Sequence 6, Appli
      9    74.5    23.0     70  12  US-10-180-247-2            Sequence 2, Appli
     10    71.5    22.1     61  15  US-10-264-480-6            Sequence 6, Appli
     11    71.5    22.1     61  15  US-10-264-480-7            Sequence 7, Appli
     12    71.5    22.1     61  15  US-10-264-480-8            Sequence 8, Appli
     13      71    21.9     47  15  US-10-178-213-411          Sequence 411, App
     14      71    21.9     77  15  US-10-178-213-410          Sequence 410, App
     15    70.5    21.8     47  15  US-10-178-213-387          Sequence 387, App
     16    70.5    21.8     65  15  US-10-264-480-14           Sequence 14, Appl
     17    70.5    21.8     84  15  US-10-178-213-386          Sequence 386, App
     18    70.5    21.8     87  15  US-10-264-480-10           Sequence 10, Appl
     19    69.5    21.5     84  15  US-10-264-480-18           Sequence 18, Appl
     20      69    21.3     47  15  US-10-178-213-408          Sequence 408, App
     21      69    21.3     47  15  US-10-178-213-444          Sequence 444, App
     22      69    21.3     77  15  US-10-178-213-407          Sequence 407, App
     23      69    21.3     78  15  US-10-178-213-443          Sequence 443, App
c    24    68.5    22.6    173  15  US-10-156-761-7787         Sequence 7787, Ap
     25      68    21.0     47  15  US-10-178-213-396          Sequence 396, App
     26      68    21.0     76  15  US-10-178-213-395          Sequence 395, App
     27    67.5    20.8     87  15  US-10-264-480-12           Sequence 12, Appl
c    28    67.5    22.3    439  15  US-10-081-872-208          Sequence 208, App
     29      67    20.7     45  15  US-10-178-213-168          Sequence 168, App
     30      67    20.7     72  15  US-10-178-213-167          Sequence 167, App
     31      67    20.7     86  15  US-10-264-480-22           Sequence 22, Appl
     32      67    20.7     86  15  US-10-264-480-23           Sequence 23, Appl
     33    66.5    20.5     47  15  US-10-178-213-309          Sequence 309, App
     34    66.5    20.5     47  15  US-10-178-213-339          Sequence 339, App
     35    66.5    20.5     58  15  US-10-264-480-4            Sequence 4, Appli
     36    66.5    20.5     62  14  US-10-044-359-21           Sequence 21, Appl
     37    66.5    20.5     66  15  US-10-264-480-15           Sequence 15, Appl
     38    66.5    20.5     77  15  US-10-178-213-308          Sequence 308, App
     39    66.5    20.5     84  15  US-10-178-213-338          Sequence 338, App
c    40    66.5    21.9    810  15  US-10-156-761-10081        Sequence 10081, A
c    41      66    21.8    198  10  US-09-764-868-677          Sequence 677, App
c    42      66    21.8    307  10  US-09-764-868-1100         Sequence 1100, Ap
c    43      66    21.8    312  9  US-09-925-302-783          Sequence 783, App
     44    65.5    20.2     44  15  US-10-178-213-147          Sequence 147, App
     45    65.5    20.2     58  15  US-10-264-480-3            Sequence 3, Appli

                                 ALIGNMENTS

RESULT 1
US-10-125-258-119
; Sequence 119, Application US/10125258
; Publication No. US20030028920A1
; GENERAL INFORMATION:
;  APPLICANT: Altier, Daniel J.
;  APPLICANT:  Herrmann, Rafael
;  APPLICANT:  Lu, Albert L.
;  APPLICANT:  McCutchen, Billy F.
;  APPLICANT:  Presnail, James K.
;  APPLICANT:  Weaver, Janine L.
;  APPLICANT:  Wong, James F. H.
;  TITLE OF INVENTION: Antimicrobial Polypeptides and Their
;  TITLE OF INVENTION:  Uses
;  FILE REFERENCE: 35718/246215
;  CURRENT APPLICATION NUMBER: US/10/125,258
;  CURRENT FILING DATE:  2002-04-18
;  PRIOR APPLICATION NUMBER: 60/285,355
;  PRIOR FILING DATE: 2001-04-20
;  NUMBER OF SEQ ID NOS: 127
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 119
;   LENGTH: 44
;   TYPE: PRT
```

```
;   ORGANISM: Agrotis ipsilon
US-10-125-258-119

Alignment Scores:
Pred. No.:               8.34e-21      Length:         44
Score:                   232.00        Matches:        38
Percent Similarity:      93.02%        Conservative:   2
Best Local Similarity:   88.37%        Mismatches:     3
Query Match:             71.60%        Indels:         0
DB:                      15            Gaps:           0

US-09-673-274B-2 (1-169) x US-10-125-258-119 (1-44)

Qy          1 GATAAGCTTATCGGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
                :       ||||| : ||||:: |||| :|||:::::: |||: |:::|||:
Db          1 AspValHisIleGlySerCysValTrpGlyAlaValAspTyrThrSerAsnCysAsnAsn 20

Qy         61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
              : ||||||:|:||||| | ||||:| |||::.|||| |||:'||| :||| ||||:||||
Db         21 GluCysLysArgArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnValAsnCys 40

Qy        121 TGGTGCGAG 129
              ||||:||||
Db         41 TrpCysGlu 43

RESULT 2
US-10-125-258-117
; Sequence 117, Application US/10125258
; Publication No. US20030028920A1
; GENERAL INFORMATION:
;  APPLICANT: Altier, Daniel J.
;  APPLICANT:  Herrmann, Rafael
;  APPLICANT:  Lu, Albert L.
;  APPLICANT:  McCutchen, Billy F.
;  APPLICANT:  Presnail, James K.
;  APPLICANT:  Weaver, Janine L.
;  APPLICANT:  Wong, James F. H.
;  TITLE OF INVENTION: Antimicrobial Polypeptides and Their
;  TITLE OF INVENTION:  Uses
;  FILE REFERENCE: 35718/246215
;  CURRENT APPLICATION NUMBER: US/10/125,258
;  CURRENT FILING DATE:  2002-04-18
;  PRIOR APPLICATION NUMBER: 60/285,355
;  PRIOR FILING DATE: 2001-04-20
;  NUMBER OF SEQ ID NOS: 127
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 117
;   LENGTH: 73
;   TYPE: PRT
;   ORGANISM: Agrotis ipsilon
US-10-125-258-117

Alignment Scores:
Pred. No.:               8.94e-21      Length:         73
Score:                   232.00        Matches:        38
Percent Similarity:      93.02%        Conservative:   2
Best Local Similarity:   88.37%        Mismatches:     3
Query Match:             71.60%        Indels:         0
DB:                      15            Gaps:           0

US-09-673-274B-2 (1-169) x US-10-125-258-117 (1-73)

Qy          1 GATAAGCTTATCGGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
              |.        ||||||||||:|:|||::|||.:|||:::||·:|||||:::||||||
Db         30 AspValHisIleGlySerCysValTrpGlyAlaValAspTyrThrSerAsnCysAsnAsn 49

Qy         61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
              | |:||||:|||||||:||||||:||||| :|||:. |||: |||:|||| ||||:||||:|
Db         50 GluCysLysArgArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnValAsnCys 69

Qy        121 TGGTGCGAG 129
              |:||||||
Db         70 TrpCysGlu 72
```

```
RESULT 3
US-10-264-480-13
; Sequence 13, Application US/10264480
; Publication No. US20030113892A1
; GENERAL INFORMATION:
;  APPLICANT: Hammock, Bruce D.
;  APPLICANT:  Inceoglu, Bora
;  TITLE OF INVENTION: ISOLATED POLYPEPTIDES AND COMPOSITIONS
;  TITLE OF INVENTION:  FROM THE VENOM OF P. TRANSVAALICUS AND METHODS OF USE
;  FILE REFERENCE: UCAL256
;  CURRENT APPLICATION NUMBER: US/10/264,480
;  CURRENT FILING DATE:  2002-10-04
;  PRIOR APPLICATION NUMBER: 60/393,070
;  PRIOR FILING DATE: 2002-06-28
;  PRIOR APPLICATION NUMBER: 60/327,602
;  PRIOR FILING DATE: 2001-10-04
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13
;   LENGTH: 65
;   TYPE: PRT
;   ORGANISM: Centruroides sculpturatus
US-10-264-480-13

Alignment Scores:
Pred. No.:               0.162         Length:         65
Score:                   79.50         Matches:        17
Percent Similarity:      51.28%        Conservative:   3
Best Local Similarity:   43.59%        Mismatches:     12
Query Match:             24.54%        Indels:         7
DB:                      15            Gaps:           3

US-09-673-274B-2 (1-169) x US-10-264-480-13 (1-65)

Qy         19 TGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGGGGT 78
               |||   |||             |||                |||||| |||||||||   :::
Db         16 CysPheTrpLeuGlyLysAsn------GluHisCysAsnThrGluCysLysAlaLysAsn 33

Qy         79 TACAAGGGT------GGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
                    '||      |.|::::|||   :::||||||            |||||||||||
Db         34 GlnGlyGlySerTyrGlyTyrCysTyrAlaPheAla---------CysTrpCysGlu 49

RESULT 4
US-10-264-480-11
; Sequence 11, Application US/10264480
; Publication No. US20030113892A1
; GENERAL INFORMATION:
;  APPLICANT: Hammock, Bruce D.
;  APPLICANT:  Inceoglu, Bora
;  TITLE OF INVENTION: ISOLATED POLYPEPTIDES AND COMPOSITIONS
;  TITLE OF INVENTION:  FROM THE VENOM OF P. TRANSVAALICUS AND METHODS OF USE
;  FILE REFERENCE: UCAL256
;  CURRENT APPLICATION NUMBER: US/10/264,480
;  CURRENT FILING DATE:  2002-10-04
;  PRIOR APPLICATION NUMBER: 60/393,070
;  PRIOR FILING DATE: 2002-06-28
;  PRIOR APPLICATION NUMBER: 60/327,602
;  PRIOR FILING DATE: 2001-10-04
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
;   LENGTH: 66
;   TYPE: PRT
;   ORGANISM: Centruroides limpidus limpidus
US-10-264-480-11

Alignment Scores:
Pred. No.:               0.162         Length:         66
Score:                   79.50         Matches:        17
Percent Similarity:      53.85%        Conservative:   4
Best Local Similarity:   43.59%        Mismatches:     11
Query Match:             24.54%        Indels:         7
```

```
B:                       15            Gaps:           3

S-09-673-274B-2 (1-169) x US-10-264-480-11 (1-66)

y          19 TGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGGGGT 78
              |||   |||         |||         :::|||:::   |||||:|||   :::
b          16 CysPheTrpLeuGlyLysAsn------GluAsnCysAspLysGluCysLysAlaLysAsn 33

y          79 TACAAGGGT------GGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
                    |||     |||:::||| |||||||||       ||:||||:|||
b          34 GlnGlyGlySerTyrGlyTyrCysTyrSerPheAla--------CysTrpCysGlu 49

ESULT 5
S-10-264-480-9
 Sequence 9, Application US/10264480
 Publication No. US20030113892A1
 GENERAL INFORMATION:
  APPLICANT: Hammock, Bruce D.
  APPLICANT:  Inceoglu, Bora
  TITLE OF INVENTION: ISOLATED POLYPEPTIDES AND COMPOSITIONS
  TITLE OF INVENTION:  FROM THE VENOM OF P. TRANSVAALICUS AND METHODS OF USE
  FILE REFERENCE: UCAL256
  CURRENT APPLICATION NUMBER: US/10/264,480
  CURRENT FILING DATE:  2002-10-04
  PRIOR APPLICATION NUMBER: 60/393,070
  PRIOR FILING DATE: 2002-06-28
  PRIOR APPLICATION NUMBER: 60/327,602
  PRIOR FILING DATE: 2001-10-04
  NUMBER OF SEQ ID NOS: 30
  SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 9
   LENGTH: 61
   TYPE: PRT
   ORGANISM: Leiurus quinquestriatus quinquestriatus
S-10-264-480-9

lignment Scores:
red. No.:                0.384         Length:         61
core:                    76.50         Matches:        15
ercent Similarity:       55.26%        Conservative:   6
est Local Similarity:    39.47%        Mismatches:     12
uery Match:              23.61%        Indels:         5
B:                       15            Gaps:           2

S-09-673-274B-2 (1-169) x US-10-264-480-9 (1-61)

y          16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
              ||||||::::::|||                     |||:||  ||||||:||
b          13 SerCysLeuPheGly------------AsnGluGlyCysAsnLysGluCysLysSerTyr 28

y          76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
              |||         |||:::|||   ::::::      :::   ||||:|||:|||
b          29 GlyGlySerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 45

ESULT 6
S-09-829-481-2
 Sequence 2, Application US/09829481
 Patent No. US20020069427A1
, GENERAL INFORMATION:
,  APPLICANT: Presnail, James
,  APPLICANT:  Weng, Zude
,  APPLICANT:  Wong, James
,  TITLE OF INVENTION: Arthropod Defensins
,  FILE REFERENCE: BB1441 US NA
,  CURRENT APPLICATION NUMBER: US/09/829,481
,  CURRENT FILING DATE:  2001-04-10
,  PRIOR APPLICATION NUMBER: 60/197279
,  PRIOR FILING DATE: 2000-04-14
,  NUMBER OF SEQ ID NOS: 11
,  SOFTWARE: Microsoft Office 97
, SEQ ID NO 2
,   LENGTH: 61
,   TYPE: PRT
```

```
;   ORGANISM: Scolopendra canidens DS
US-09-829-481-2

Alignment Scores:
Pred. No.:                0.594           Length:         61
Score:                    75.00           Matches:        11
Percent Similarity:       53.57%          Conservative:   4
Best Local Similarity:    39.29%          Mismatches:     13
Query Match:              23.15%          Indels:         0
DB:                       9               Gaps:           0

US-09-673-274B-2 (1-169) x US-09-829 481 2 (1-61)

Qy          49 GATTGCAACGGTGAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCT 108
               :::||||||       ||:::       ||:::   |||:|.:::|.|      :.|
Db          34 GluCysAsnArgHisCysArgGlyAsnGlyPheThrGlyGlyTyrCysThrGlyPheLeu 53

Qy         109 AACGTGAACTGCTGGTGCGAGACT 132
                        | |   |:|   ::|
Db          54 LysPheThrCysLysCysTyrThr 61

RESULT 7
US-10-180-247-4
; Sequence 4, Application US/10180247
; Publication No. US20030167519A1
; GENERAL INFORMATION:
;  APPLICANT: DeRose, Richard
;  APPLICANT:  Freysinnet, Georges
;  APPLICANT:  Hoffman, Jules
;  TITLE OF INVENTION: Chimeric Gene Encoding Drosomycin,
;  TITLE OF INVENTION:  Vector Containing It and Production of Disease-Resistant
;  TITLE OF INVENTION:  Transgenic Plants
;  FILE REFERENCE: A32889-PCT-USA A-A 072667.0182
;  CURRENT APPLICATION NUMBER: US/10/180,247
;  CURRENT FILING DATE:  2002-06-26
;  PRIOR APPLICATION NUMBER: 09/480,251
;  PRIOR FILING DATE: 2000-01-11
;  PRIOR APPLICATION NUMBER: PCT/FR98/01462
;  PRIOR FILING DATE: 1998-07-08
;  PRIOR APPLICATION NUMBER: FR97/09,115
;  PRIOR FILING DATE: 1997-07-11
;  PRIOR APPLICATION NUMBER: FR97/09,663
;  PRIOR FILING DATE: 1997-07-24
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 44
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-10-180-247-4

Alignment Scores:
Pred. No.:                0.657           Length:         44
Score:                    74.50           Matches:        15
Percent Similarity:       42.50%          Conservative:   2
Best Local Similarity:    37.50%          Mismatches:     16
Query Match:              22.99%          Indels:         7
DB:                       12              Gaps:           3

US-09-673-274B-2 (1-169) x US-10-180-247-4 (1-44)

Qy          13 GGTTCCTGC---GTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAG 69
               |||   |||   |||:||                        |||         |||||:
Db           9 GlyProCysAlaValTrpAsp---------- --AsnGluThrCysArgArgValCysLys 24

Qy          70 AGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
                     ::|         |||||: |||            ::::::   |||||:| ||||
Db          25 GluGluGlyArgSerSerGlyHisCysSer------ProSerLeuLysCysTrpCysGlu 42

RESULT 8
US-10-180-247-6
; Sequence 6, Application US/10180247
; Publication No. US20030167519A1
```

```
; GENERAL INFORMATION:
;  APPLICANT: DeRose, Richard
;  APPLICANT:  Freysinnet, Georges
;  APPLICANT:  Hoffman, Jules
;  TITLE OF INVENTION: Chimeric Gene Encoding Drosomycin,
;  TITLE OF INVENTION:  Vector Containing It and Production of Disease-Resistant
;  TITLE OF INVENTION:  Transgenic Plants
;  FILE REFERENCE: A32889-PCT-USA-A-A 072667.0182
;  CURRENT APPLICATION NUMBER: US/10/180,247
;  CURRENT FILING DATE:  2002-06-26
;  PRIOR APPLICATION NUMBER: 09/480,251
;  PRIOR FILING DATE: 2000-01-11
;  PRIOR APPLICATION NUMBER: PCT/FR98/01462
;  PRIOR FILING DATE: 1998-07-08
;  PRIOR APPLICATION NUMBER: FR97/09,115
;  PRIOR FILING DATE: 1997-07-11
;  PRIOR APPLICATION NUMBER: FR97/09,663
;  PRIOR FILING DATE: 1997-07-24
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 69
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Ubiquitin drosomycin fusion peptide
US-10-180-247-6

Alignment Scores:
Pred. No.:                0.699          Length:        69
Score:                    74.50          Matches:       15
Percent Similarity:       42.50%         Conservative:  2
Best Local Similarity:    37.50%         Mismatches:    16
Query Match:              22.99%         Indels:        7
DB:                       12             Gaps:          3

US-09-673-274B-2 (1-169) x US-10-180-247-6 (1-69)

Qy         13 GGTTCCTGC---GTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAG 69
              .:    ||.   ||||::                      |||        ||::||
Db         34 GlyProCysAlaValTrpAsp ------------AsnGluThrCysArgArgValCysLys 49

Qy         70 AGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
                    |||         |.||||.:           ::::::   |||:|||:|||
Db         50 GluGluGlyArgSerSerGlyHisCysSer------ProSerLeuLysCysTrpCysGlu 67

RESULT 9
US-10-180-247-2
; Sequence 2, Application US/10180247
; Publication No. US20030167519A1
; GENERAL INFORMATION:
;  APPLICANT: DeRose, Richard
;  APPLICANT:  Freysinnet, Georges
;  APPLICANT:  Hoffman, Jules
;  TITLE OF INVENTION: Chimeric Gene Encoding Drosomycin,
;  TITLE OF INVENTION:  Vector Containing It and Production of Disease-Resistant
;  TITLE OF INVENTION:  Transgenic Plants
;  FILE REFERENCE: A32889-PCT-USA-A-A 072667.0182
;  CURRENT APPLICATION NUMBER: US/10/180,247
;  CURRENT FILING DATE:  2002-06-26
;  PRIOR APPLICATION NUMBER: 09/480,251
;  PRIOR FILING DATE: 2000-01-11
;  PRIOR APPLICATION NUMBER: PCT/FR98/01462
;  PRIOR FILING DATE: 1998-07-08
;  PRIOR APPLICATION NUMBER: FR97/09,115
;  PRIOR FILING DATE: 1997-07-11
;  PRIOR APPLICATION NUMBER: FR97/09,663
;  PRIOR FILING DATE: 1997-07-24
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 70
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-10-180-247-2

Alignment Scores:
Pred. No.:                0.7            Length:        70
Score:                    74.50          Matches:       15
Percent Similarity:       42.50%         Conservative:  2
Best Local Similarity:    37.50%         Mismatches:    16
Query Match:              22.99%         Indels:        7
DB:                       12             Gaps:          3

US-09-673-274B-2 (1-169) x US-10-180-247-2 (1-70)

Qy         13 GGTTCCTGC---GTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAG 69
              |.|   .||   |||:.|                      |||        ||:|:|
Db         35 GlyProCysAlaValTrpAsp------------AsnGluThrCysArgArgValCysLys 50

Qy         70 AGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
                    |||         ||||:||||           ::::::   ||||:|||||||
Db         51 GluGluGlyArgSerSerGlyHisCysSer------ProSerLeuLysCysTrpCysGlu 68

RESULT 10
US-10-264-480-6
; Sequence 6, Application US/10264480
; Publication No. US20030113892A1
; GENERAL INFORMATION:
;  APPLICANT: Hammock, Bruce D.
;  APPLICANT:  Inceoglu, Bora
;  TITLE OF INVENTION: ISOLATED POLYPEPTIDES AND COMPOSITIONS
;  TITLE OF INVENTION:  FROM THE VENOM OF P. TRANSVAALICUS AND METHODS OF USE
;  FILE REFERENCE: UCAL256
;  CURRENT APPLICATION NUMBER: US/10/264,480
;  CURRENT FILING DATE:  2002-10-04
;  PRIOR APPLICATION NUMBER: 60/393,070
;  PRIOR FILING DATE: 2002-06-28
;  PRIOR APPLICATION NUMBER: 60/327,602
;  PRIOR FILING DATE: 2001-10-04
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 61
;   TYPE: PRT
;   ORGANISM: Buthacus arenicola
US-10-264-480-6

Alignment Scores:
Pred. No.:                1.65           Length:        61
Score:                    71.50          Matches:       14
Percent Similarity:       55.26%         Conservative:  7
Best Local Similarity:    36.84%         Mismatches:    12
Query Match:              22.07%         Indels:        5
DB:                       15             Gaps:          2

US-09-673-274B-2 (1-169) x US-10-264-480-6 (1-61)

Qy         16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
              |:||||::::::||:                     |||:::   |||||||||
Db         13 SerCysLeuPheGly------------AsnGluGlyCysAspLysGluCysLysAlaTyr 28

Qy         76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
              |:|         ||:::|||    ::::::      :::   ||||||||||||
Db         29 GlyGlySerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 45

RESULT 11
US-10-264-480-7
; Sequence 7, Application US/10264480
; Publication No. US20030113892A1
; GENERAL INFORMATION:
;  APPLICANT: Hammock, Bruce D.
;  APPLICANT:  Inceoglu, Bora
;  TITLE OF INVENTION: ISOLATED POLYPEPTIDES AND COMPOSITIONS
;  TITLE OF INVENTION:  FROM THE VENOM OF P. TRANSVAALICUS AND METHODS OF USE
;  FILE REFERENCE: UCAL256
```

```
  CURRENT APPLICATION NUMBER: US/10/264,480
  CURRENT FILING DATE:  2002-10-04
  PRIOR APPLICATION NUMBER: 60/393,070
  PRIOR FILING DATE: 2002-06-28
  PRIOR APPLICATION NUMBER: 60/327,602
  PRIOR FILING DATE: 2001-10-04
  NUMBER OF SEQ ID NOS: 30
  SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 7
   LENGTH: 61
   TYPE: PRT
   ORGANISM: Buthus occitanus
US-10-264-480-7

Alignment Scores:
Pred. No.:              1.65           Length:         61
Score:                  71.50          Matches:        14
Percent Similarity:     55.26%         Conservative:   7
Best Local Similarity:  36.84%         Mismatches:     12
Query Match:            22.07%         Indels:         5
DB:                     15             Gaps:           2

US-09-673-274B-2 (1-169) x US-10-264-480-7 (1-61)

Qy          16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
               ||||||::::::|||                     |||:::   |||||||||
Db          13 SerCysLeuPheGly------------AsnGluGlyCysAspLysGluCysLysAlaTyr 28

Qy          76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
               |||         |||:::|||   ::::::      :::   ||||||||||||
Db          29 GlyGlySerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 45

RESULT 12
US-10-264-480-8
; Sequence 8, Application US/10264480
; Publication No. US20030113892A1
; GENERAL INFORMATION:
;  APPLICANT: Hammock, Bruce D.
;  APPLICANT:  Inceoglu, Bora
;  TITLE OF INVENTION: ISOLATED POLYPEPTIDES AND COMPOSITIONS
   TITLE OF INVENTION:  FROM THE VENOM OF P. TRANSVAALICUS AND METHODS OF USE
;  FILE REFERENCE: UCAL256
;  CURRENT APPLICATION NUMBER: US/10/264,480
;  CURRENT FILING DATE:  2002-10-04
;  PRIOR APPLICATION NUMBER: 60/393,070
;  PRIOR FILING DATE: 2002-06-28
;  PRIOR APPLICATION NUMBER: 60/327,602
;  PRIOR FILING DATE: 2001-10-04
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
;   LENGTH: 61
;   TYPE: PRT
;   ORGANISM: Buthus occitanus
US-10-264-480-8

Alignment Scores:
Pred. No.:              1.65           Length:         61
Score:                  71.50          Matches:        14
Percent Similarity:     55.26%         Conservative:   7
Best Local Similarity:  36.84%         Mismatches:     12
Query Match:            22.07%         Indels:         5
DB:                     15             Gaps:           2

US-09-673-274B-2 (1-169) x US-10-264-480-8 (1-61)

Qy          16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
               ||||||::::::|||                     |||:::   |||||||||
Db          13 SerCysLeuPheGly------------AsnGluGlyCysAspLysGluCysLysAlaTyr 28

Qy          76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
               |||         |||:::|||   ::::::      :::   ||||||||||||
Db          29 GlyGlySerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 45
```

```
RESULT 13
US-10-178-213-411
; Sequence 411, Application US/10178213
; Publication No. US20030041348A1
; GENERAL INFORMATION:
;  APPLICANT: Simmons, Carl R.
;  APPLICANT:  Navarro Acevedo, Pedro A.
;  APPLICANT:  Harvell, Leslie
;  APPLICANT:  Cahoon, Rebecca
;  APPLICANT:  McCutchen, Billy Fred
;  APPLICANT:  Lu, Albert
;  APPLICANT:  Herrmann, Rafael
;  APPLICANT:  Wong, James
;  TITLE OF INVENTION: Defensin Polynucleotides and Methods of
;  TITLE OF INVENTION:  Use
;  FILE REFERENCE: 35718/246703
;  CURRENT APPLICATION NUMBER: US/10/178,213
;  CURRENT FILING DATE:  2002-06-21
;  PRIOR APPLICATION NUMBER: 60/300,152
;  PRIOR FILING DATE: 2001-06-22
;  PRIOR APPLICATION NUMBER: 60/300,241
;  PRIOR FILING DATE: 2001-06-22
;  NUMBER OF SEQ ID NOS: 469
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 411
;   LENGTH: 47
;   TYPE: PRT
;   ORGANISM: Tropaeolum majus
US-10-178-213-411

Alignment Scores:
Pred. No.:              1.84           Length:         47
Score:                  71.00          Matches:        14
Percent Similarity:     50.00%         Conservative:   5
Best Local Similarity:  36.84%         Mismatches:     13
Query Match:            21.91%         Indels:         6
DB:                     15             Gaps:           2

US-09-673-274B-2 (1-169) x US-10-178-213-411 (1-47)

Qy          13 GGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGG 72
               |||   |.:|||                     ::::::|||          :||||
Db          12 GlyProCysVal---------------SerLysThrAsnCysAlaSerValCysLysThr 26

Qy          73 AGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGC 126
                  |||:::   |:||||  :|||      |||   :::      |||:::|||
Db          27 GluGlyPheTyrGlyGlyHisCysArgGlyPheArgHis---ArgCysPheCys 43

RESULT 14
US-10-178-213-410
; Sequence 410, Application US/10178213
; Publication No. US20030041348A1
; GENERAL INFORMATION:
;  APPLICANT: Simmons, Carl R.
;  APPLICANT:  Navarro Acevedo, Pedro A.
;  APPLICANT:  Harvell, Leslie
;  APPLICANT:  Cahoon, Rebecca
;  APPLICANT:  McCutchen, Billy Fred
;  APPLICANT:  Lu, Albert
;  APPLICANT:  Herrmann, Rafael
;  APPLICANT:  Wong, James
;  TITLE OF INVENTION: Defensin Polynucleotides and Methods of
;  TITLE OF INVENTION:  Use
;  FILE REFERENCE: 35718/246703
;  CURRENT APPLICATION NUMBER: US/10/178,213
;  CURRENT FILING DATE:  2002-06-21
;  PRIOR APPLICATION NUMBER: 60/300,152
;  PRIOR FILING DATE: 2001-06-22
;  PRIOR APPLICATION NUMBER: 60/300,241
;  PRIOR FILING DATE: 2001-06-22
;  NUMBER OF SEQ ID NOS: 469
;  SOFTWARE: FastSEQ for Windows Version 4.0
```

```
; SEQ ID NO 410
;   LENGTH: 77
;   TYPE: PRT
;   ORGANISM: Tropaeolum majus
US-10-178-213-410

Alignment Scores:
Pred. No.:               1.96          Length:        77
Score:                   71.00         Matches:       14
Percent Similarity:      50.00%        Conservative:  5
Best Local Similarity:   36.84%        Mismatches:    13
Query Match:             21.91%        Indels:        6
DB:                      15            Gaps:          2

US-09-673-274B-2 (1-169) x US-10-178-213-410 (1-77)

Qy          13 GGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGG 72
               |||    ||||:                     :::::::|||         ||||||
Db          42 GlyProCysVal---------------SerLysThrAsnCysAlaSerValCysLysThr 56

Qy          73 AGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGC 126
                  .||:::   ||||.:|||||       .||    :::      '|::: ||
Db          57 GluGlyPheTyrGlyGlyHisCysArgGlyPheArgHis---ArgCysPheCys 73

RESULT 15
US-10-178-213-387
; Sequence 387, Application US/10178213
; Publication No. US20030041348A1
; GENERAL INFORMATION:
;  APPLICANT: Simmons, Carl R.
;  APPLICANT:  Navarro Acevedo, Pedro A.
;  APPLICANT:  Harvell, Leslie
;  APPLICANT:  Cahoon, Rebecca
;  APPLICANT:  McCutchen, Billy Fred
;  APPLICANT:  Lu, Albert
;  APPLICANT:  Herrmann, Rafael
;  APPLICANT:  Wong, James
;  TITLE OF INVENTION: Defensin Polynucleotides and Methods of
;  TITLE OF INVENTION:  Use
;  FILE REFERENCE: 35718/246703
;  CURRENT APPLICATION NUMBER: US/10/178,213
;  CURRENT FILING DATE:  2002-06-21
;  PRIOR APPLICATION NUMBER: 60/300,152
;  PRIOR FILING DATE: 2001-06-22
;  PRIOR APPLICATION NUMBER: 60/300,241
;  PRIOR FILING DATE: 2001-06-22
;  NUMBER OF SEQ ID NOS: 469
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 387
;   LENGTH: 47
;   TYPE: PRT
;   ORGANISM: Glycine max
US-10-178-213-387

Alignment Scores:
Pred. No.:               2.12          Length:        47
Score:                   70.50         Matches:       10
Percent Similarity:      59.26%        Conservative:  6
Best Local Similarity:   37.04%        Mismatches:    10
Query Match:             21.76%        Indels:        1
DB:                      15            Gaps:          1

US-09-673-274B-2 (1-169) x US-10-178-213-387 (1-47)

Qy          46 TCCGATTGCAACGGTGAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTC 105
               ::::::|||         |||:::      |||:::   .|||:.|||:||      |||
Db          18 ThrAsnCysGlySerValCysArgThrGluGlyPheSerGlyGlyHisCysArgGlyPhe 37

Qy         106 GCTAACGTGAACTGCTGGTGC 126
               :::         |||:::|||
Db          38 SerArg---ArgCysPheCys 43
```

```
Search completed: October 16, 2003, 17:34:07
Job time : 36.0914 secs
```

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2003  Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on:        October 16, 2003, 17:06:37 ; Search time 16.8439 Seconds
                                             (without alignments)
                                             849.038 Million cell updates/sec

Title:         US-09-673-274B-2
Perfect score: 324
Sequence:      1 gataagcttatcggttcctg..........gaacgtgtcgacggatccgg 169

Scoring table: BLOSUM62
               Xgapop 10.0 , Xgapext  0.5
               Ygapop 10.0 , Ygapext  0.5
               Fgapop  6.0 , Fgapext  7.0
               Delop   6.0 , Delext   7.0

Searched:      328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters:      657434

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09673274/runat_16102003_170628_2580/app_query.fasta_1.654
-DB=Issued_Patents_AA -QFMT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09673274_@CGN_1_1_55_@runat_16102003_170628_2580 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :     Issued_Patents_AA:*
               1:  /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
               2:  /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
               3:  /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
               4:  /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
               5:  /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
               6:  /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                SUMMARIES
               %
Result         Query
  No.   Score  Match Length DB  ID                         Description
-----------------------------------------------------------------------------
c  1     81.5   26.9    280  4  US-09-252-991A-19400        Sequence 19400, A
   2     80.5   24.8     65  1  US-08-435-040-2             Sequence 2, Appli
   3     80.5   24.8     65  3  US-09-020-216-2             Sequence 2, Appli
c  4       78   25.7    149  4  US-09-252-991A-32270        Sequence 32270, A
   5     76.5   23.6     61  1  US-08-682-485A-25           Sequence 25, Appl
   6     76.5   23.6     61  1  US-08-451-472-4             Sequence 4, Appli
   7     76.5   23.6     61  2  US-08-933-314-25            Sequence 25, Appl
   8       75   23.1     64  4  US-09-403-343B-25           Sequence 25, Appl
   9     74.5   23.0     44  4  US-09-480-251-4             Sequence 4, Appli
  10     74.5   23.0     69  4  US-09-480-251-6             Sequence 6, Appli
  11     74.5   23.0     70  4  US-09-480-251-2             Sequence 2, Appli
c 12     73.5   24.3    158  4  US-09-252-991A-26807        Sequence 26807, A
c 13       73   24.1    145  4  US-09-252-991A-22334        Sequence 22334, A
c 14     71.5   23.6    248  4  US-09-252-991A-17266        Sequence 17266, A
c 15       70   23.1    397  4  US-09-252-991A-26857        Sequence 26857, A
c 16     68.5   22.6    147  4  US-09-252-991A-16810        Sequence 16810, A
c 17       68   22.4    296  4  US-09-252-991A-32162        Sequence 32162, A
  18       67   20.7     60  1  US-08-451-472-6             Sequence 6, Appli
  19       67   20.7     67  1  US-08-435-040-3             Sequence 3, Appli
  20       67   20.7     67  3  US-09-020-216-3             Sequence 3, Appli
c 21       67   22.1    355  4  US-09-252-991A-23003        Sequence 23003, A
c 22       67   22.1    631  4  US-09-252-991A-20063        Sequence 20063, A
c 23     66.5   21.9    167  4  US-09-252-991A-31970        Sequence 31970, A
c 24     66.5   21.9    187  4  US-09-252-991A-27028        Sequence 27028, A
c 25       66   21.8    437  4  US-09-252-991A-20209        Sequence 20209, A
c 26       66   21.8    781  4  US-09-252-991A-26207        Sequence 26207, A
c 27     65.5   21.6    195  4  US-09-252-991A-28844        Sequence 28844, A
  28     65.5   20.2    276  4  US-09-252-991A-18128        Sequence 18128, A
c 29       65   21.5    160  4  US-09-252-991A-19367        Sequence 19367, A
  30     64.5   19.9     80  3  US-08-952-383A-14           Sequence 14, Appl
  31     64.5   19.9     80  4  US-08-970-264A-27           Sequence 27, Appl
c 32     64.5   21.3    169  4  US-09-252-991A-22731        Sequence 22731, A
c 33     64.5   21.3    187  4  US-09-252-991A-19143        Sequence 19143, A
  34     63.5   19.6     61  1  US-08-682-485A-26           Sequence 26, Appl
  35     63.5   19.6     61  2  US-08-933-314-26            Sequence 26, Appl
c 36     63.5   21.0    246  4  US-09-252-991A-16586        Sequence 16586, A
c 37     63.5   21.0    354  4  US-09-252-991A-19736        Sequence 19736, A
c 38     63.5   21.0    399  4  US-09-252-991A-18432        Sequence 18432, A
c 39     63.5   21.0    510  4  US-09-252-991A-19528        Sequence 19528, A
c 40       63   20.8    270  4  US-09-252-991A-17589        Sequence 17589, A
c 41     62.5   20.6    138  4  US-09-252-991A-31377        Sequence 31377, A
c 42     62.5   20.6    138  4  US-09-252-991A-32675        Sequence 32675, A
c 43     62.5   20.6    641  4  US-09-252-991A-26994        Sequence 26994, A
c 44       62   20.5    134  4  US-09-252-991A-20796        Sequence 20796, A
c 45       62   20.5    141  4  US-09-252-991A-23611        Sequence 23611, A


                               ALIGNMENTS

RESULT 1
US-09-252-991A-19400
; Sequence 19400, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 19400
;   LENGTH: 280
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19400

Alignment Scores:
Pred. No.:                 0.0241          Length:         280
Score:                     81.50           Matches:        22
Percent Similarity:        49.09%          Conservative:   5
Best Local Similarity:     40.00%          Mismatches:     19
Query Match:               26.90%          Indels:         9
DB:                        4               Gaps:           3

US-09-673-274B-2 (1-169) x US-09-252-991A-19400 (1-280)

Qy        160 TCGACACGTTCGCCTCGCCGAGCTCTCAAGTCTCGCACCAGCAGTTCACGTTAGCGAAGG 101
              |||   ||||||:||||||'|||        :::   ||,|||::::::   |||   :||
Db        180 SerIleArgSerProArgArgArgThrArgProArgThrAlaAlaGlyArgGlyArgHis 199
```

```
Qy        100 AACCGCAGTGACCACCCTTGT---------AACCCCTCCTCTTGC------------ACT 62
                 |||:::         :||            |||      ..|
Db        200 ArgArgAla------AlaCysArgGlyGlyLysProArgAspCysGlyCysArgTrpCys 217

Qy         61 CACCGTTGCAATCGGAAGTGTAGTTCACAGCACCCCACACGCAGG 17
              |||||:|||   .||      |||:::      |||   |||. |
Db        218 HisArgCysArgArgProGlySerThrAlaGlyProAlaArgArg 232

RESULT 2
US-08-435-040-2
; Sequence 2, Application US/08435040
; Patent No. 5756340
;  GENERAL INFORMATION:
;    APPLICANT:  Hammock, Bruce D.
;    APPLICANT:  Herrmann, Rafael
;    APPLICANT:  Moskowitz, Haim
;    TITLE OF INVENTION:  Insect Control With Multiple Toxins
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Majestic, Parsons, Siebert & Hsue
;      STREET:  Four Embarcadero Center, Suite 1450
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  94111-4121
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/435,040
;      FILING DATE:  08-MAY-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Siebert, J. Suzanne
;      REGISTRATION NUMBER:  28,758
;      REFERENCE/DOCKET NUMBER:  2500.078US0
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 362-5556
;      TELEFAX:  (415) 362-5418
;      TELEX:  278638 MGPS
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  65 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
US-08-435-040-2

Alignment Scores:
Pred. No.:                0.0254        Length:         65
Score:                    80.50         Matches:        17
Percent Similarity:       57.14%        Conservative:   7
Best Local Similarity:    40.48%        Mismatches:     9
Query Match:              24.85%        Indels:         9
DB:                       1             Gaps:           4

US-09-673-274B-2 (1-169) x US-08-435-040-2 (1-65)

Qy         16 TCCTGCGTGTGG------GGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAG 69
              :::||||||:::      ||||||   :::||         |||.||   ||||||
Db         12 AsnCysValTyrThrCysGlyAlaAsnSerTyr---------CysAsnThrGluCysThr 28

Qy         70 AGGAGGGGTTACAAGGGTGGTCACTGC---------GGTTCCTTCGCTAACGTGAACTGC 120
              :::   :||   :::   ||:::|||         |||   :::   |||      ||:
Db         29 LysAsnGlyAlaGluSerGlyTyrCysGlnTrpPheGlyLysTyrGlyAsnAla---Cys 47

Qy        121 TGGTGC 126
              ||||:|
Db         48 TrpCys 49

RESULT 3
US-09-020-216-2
; Sequence 2, Application US/09020216
; Patent No. 6162430
;  GENERAL INFORMATION:
;    APPLICANT:  Hammock, Bruce D.
;    APPLICANT:  Herrmann, Rafael
;    APPLICANT:  Moskowitz, Haim
;    TITLE OF INVENTION:  Insect Control With Multiple Toxins
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Majestic, Parsons, Siebert & Hsue
;      STREET:  Four Embarcadero Center, Suite 1450
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  94111-4121
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/020,216
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/435,040
;      FILING DATE:  08-MAY-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Siebert, J. Suzanne
;      REGISTRATION NUMBER:  28,758
;      REFERENCE/DOCKET NUMBER:  2500.078US0
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 362-5556
;      TELEFAX:  (415) 362-5418
;      TELEX:  278638 MGPS
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  65 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
US-09-020-216-2

Alignment Scores:
Pred. No.:                0.0254        Length:         65
Score:                    80.50         Matches:        17
Percent Similarity:       57.14%        Conservative:   7
Best Local Similarity:    40.48%        Mismatches:     9
Query Match:              24.85%        Indels:         9
DB:                       3             Gaps:           4

US-09-673-274B-2 (1-169) x US-09-020-216-2 (1-65)

Qy         16 TCCTGCGTGTGG------GGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAG 69
              :::||||||:::      ||||||   :::|||         ||||||   ||||||
Db         12 AsnCysValTyrThrCysGlyAlaAsnSerTyr---------CysAsnThrGluCysThr 28

Qy         70 AGGAGGGGTTACAAGGGTGGTCACTGC---------GGTTCCTTCGCTAACGTGAACTGC 120
              :::   |||   :::   |||:::|||         |||   :::   |||      |||
Db         29 LysAsnGlyAlaGluSerGlyTyrCysGlnTrpPheGlyLysTyrGlyAsnAla---Cys 47

Qy        121 TGGTGC 126
              ::||||
Db         48 TrpCys 49

RESULT 4
```

```
S-09-252-991A-32270
 Sequence 32270, Application US/09252991A
 Patent No. 6551795
 GENERAL INFORMATION:
  APPLICANT:  Marc J. Rubenfield et al.
  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
  FILE REFERENCE:  107196.136
  CURRENT APPLICATION NUMBER: US/09/252,991A
  CURRENT FILING DATE:  1999-02-18
  PRIOR APPLICATION NUMBER:  US 60/074,788
  PRIOR FILING DATE:  1998-02-18
  PRIOR APPLICATION NUMBER:  US 60/094,190
  PRIOR FILING DATE:  1998-07-27
  NUMBER OF SEQ ID NOS:  33142
 SEQ ID NO 32270
   LENGTH: 149
   TYPE: PRT
   ORGANISM: Pseudomonas aeruginosa
S-09-252-991A-32270

lignment Scores:
red. No.:               0.0607         Length:        149
core:                   78.00          Matches:       18
ercent Similarity:      49.06%         Conservative:  8
est Local Similarity:   33.96%         Mismatches:    21
uery Match:             25.74%         Indels:        6
B:                      4              Gaps:          2

S-09-673-274B-2 (1-169) x US-09-252-991A-32270 (1-149)

y        169 CCGGATCCGTCGACACGTTCGCCTCGCCGAGCTCTCAAGTCTCGCACCAGCAGTTCACGT 110
             |||   |||::::::|||      ||||||         ||||||:::   :::
b         14 ProSerProAlaSerArgArgArgArgArgCysSerProSerArgSerProThrGly--- 32

y        109 TAGCGAAGGAACCGCAGTGACCACCCTTGTAACCCCTCCTCTTGCACTCACCGTTGCAAT 50
                |||      |||:::         |||            |||      ||||||
b         33 ---ArgTyrThrArgAlaProTrpArgCys------------CysArgArgArgCysArg 47

y         49 CGGAAGTGTAGTTCACAGCACCCCACACGCAGGAACCGA 11
             |:|:::|||      :::   ||||||   :::   |||
b         48 ArgArgCysArgProArgArgProThrSerLysAlaArg 60

ESULT 5
S-08-682-485A-25
  Sequence 25, Application US/08682485A
  Patent No. 5763568
   GENERAL INFORMATION:
     APPLICANT:  ATKINSON, RONALD K
     APPLICANT:  HOWDEN, MERLIN E.H.
     APPLICANT:  TYLER, MARGARET I
     APPLICANT:  VONARX, EDWARD J
     TITLE OF INVENTION:  Insecticidal Toxins Derived From
     TITLE OF INVENTION:  Funnel Web (Atrax or Hadronyche Spiders)
     NUMBER OF SEQUENCES:  26
     CORRESPONDENCE ADDRESS:
       ADDRESSEE:  Zeneca, Inc.
       STREET:  1200 South 47th Street
       CITY:  Richmond
       STATE:  California
       COUNTRY:  USA
       ZIP:  94804
     COMPUTER READABLE FORM:
       MEDIUM TYPE:  Floppy disk
       COMPUTER:  IBM PC compatible
       OPERATING SYSTEM:  PC-DOS/MS-DOS
       SOFTWARE:  PatentIn Release #1.0, Version #1.25
     CURRENT APPLICATION DATA:
       APPLICATION NUMBER:  US/08/682,485A
       FILING DATE:
       CLASSIFICATION:  424
     PRIOR APPLICATION DATA:
       APPLICATION NUMBER:  US/08/682,485
```

```
;       FILING DATE:  17-JULY-1996
;       APPLICATION NUMBER:  US/08/256,933
;       FILING DATE:  27-JULY-1994
;       APPLICATION NUMBER:  WO 93/15108
;       FILING DATE:  29-JAN-1993
;       APPLICATION NUMBER:  AU PL0722
;       FILING DATE:  31-JAN-1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Shaw, Melissa A.
;       REGISTRATION NUMBER:  38,301
;       REFERENCE/DOCKET NUMBER:  FPD 5099/D1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  510-231-1542
;       TELEFAX:  510-231-1112
;   INFORMATION FOR SEQ ID NO:  25:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  61 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
;     ORIGINAL SOURCE:
;       ORGANISM:  Leiurus quinquestriatus quinquestriatus
US-08-682-485A-25

Alignment Scores:
Pred. No.:               0.0812         Length:        61
Score:                   76.50          Matches:       15
Percent Similarity:      55.26%         Conservative:  6
Best Local Similarity:   39.47%         Mismatches:    12
Query Match:             23.61%         Indels:        5
DB:                      1              Gaps:          2

US-09-673-274B-2 (1-169) x US-08-682-485A-25 (1-61)

Qy          16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
               |||.::::::::|||                     ||||||   |||:|:|||
Db          13 SerCysLeuPheGly------------AsnGluGlyCysAsnLysGluCysLysSerTyr 28

Qy          76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
               |||         :||:::|||   ::::::      :::   :| ||||::|||
Db          29 GlyGlySerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 45

RESULT 6
US-08-451-472-4
; Sequence 4, Application US/08451472
; Patent No. 5770192
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  BIOLOGICAL CONTROL AGENTS
;    NUMBER OF SEQUENCES:  73
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Nixon & Vanderhye PC
;      STREET:  8th Floor, 1100 No. 5770192th Glebe Road
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  USA
;      ZIP:  22201-4714
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/451,472
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/117,125
;      FILING DATE:  24-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mary J Wilson
```

```
;      REGISTRATION NUMBER:  32,955
;      REFERENCE/DOCKET NUMBER:  117-187
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 816 4000
;      TELEFAX:  (703) 816 4100
;      TELEX:  200797 NIXN UR
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  61 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US 08-451-472-4

Alignment Scores:
Pred. No.:                0.0812        Length:          61
Score:                    76.50         Matches:         15
Percent Similarity:       55.26%        Conservative:    6
Best Local Similarity:    39.47%        Mismatches:      12
Query Match:              23.61%        Indels:          5
DB:                       1             Gaps:            2

US-09-673-274B-2 (1-169) x US-08-451-472-4 (1-61)

Qy          16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
               ||||||::::::|||                     ||..||    ||| :|||
Db          13 SerCysLeuPheGly------------AsnGluGlyCysAsnLysGluCysLysSerTyr 28

Qy          76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
               | |         |||:::|||   ::::::      :::   ||||.|||||||
Db          29 GlyGlySerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 45

RESULT 7
US 08-933-314-25
; Sequence 25, Application US/08933314
; Patent No. 5959182
;  GENERAL INFORMATION:
;    APPLICANT:  ATKINSON, RONALD K
;    APPLICANT:  HOWDEN, MERLIN E.H.
;    APPLICANT:  TYLER, MARGARET I
;    APPLICANT:  VONARX, EDWARD J
;    TITLE OF INVENTION:  Insecticidal Toxins Derived From
;    TITLE OF INVENTION:  Funnel Web (Atrax or Hadronyche Spiders)
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Zeneca, Inc.
;      STREET:  1200 South 47th Street
;      CITY:  Richmond
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/933,314
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/682,485
;      FILING DATE:  17-JULY-1996
;      APPLICATION NUMBER:  US/08/256,933
;      FILING DATE:  27-JULY-1994
;      APPLICATION NUMBER:  WO 93/15108
;      FILING DATE:  29-JAN-1993
;      APPLICATION NUMBER:  AU PL0722
;      FILING DATE:  31-JAN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Shaw, Melissa A.
;      REGISTRATION NUMBER:  38,301
;      REFERENCE/DOCKET NUMBER:  PPD 5099/D1
```

```
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  510-231-1542
;      TELEFAX:  510-231-1112
;  INFORMATION FOR SEQ ID NO:  25:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  61 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Leiurus quinquestriatus quinquestriatus
US-08-933-314-25

Alignment Scores:
Pred. No.:                0.0812        Length:          61
Score:                    76.50         Matches:         15
Percent Similarity:       55.26%        Conservative:    6
Best Local Similarity:    39.47%        Mismatches:      12
Query Match:              23.61%        Indels:          5
DB:                       2             Gaps:            2

US-09-673-274B-2 (1-169) x US-08-933-314-25 (1-61)

Qy          16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
               |||.||::::::|.|                     ||||||   ||||:||||
Db          13 SerCysLeuPheGly------------AsnGluGlyCysAsnLysGluCysLysSerTyr 28

Qy          76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
               |||         |||:::|||   ::::::      :::   ||||||||||||
Db          29 GlyGlySerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 45

RESULT 8
US-09-403-343B-25
; Sequence 25, Application US/09403343B
; Patent No. 6555091
; GENERAL INFORMATION:
;  APPLICANT: JOLIVET-REYNAUD, COLETTE
;  APPLICANT:  PERRON, HERVE
;  APPLICANT:  MANDRAND, BERNARD
;  TITLE OF INVENTION: POLYPEPTIDE CAPABLE OF REACTING WITH ANTIBODIES OF
;  TITLE OF INVENTION:  PATIENTS SUFFERING FROM MULTIPLE SCLEROSIS AND USES
;  FILE REFERENCE: 104574
;  CURRENT APPLICATION NUMBER: US/09/403,343B
;  CURRENT FILING DATE:  1999-10-18
;  PRIOR APPLICATION NUMBER: FR/97/05679
;  PRIOR FILING DATE: 1997-04-29
;  PRIOR APPLICATION NUMBER: FR/97/16870
;  PRIOR FILING DATE: 1997-12-31
;  PRIOR APPLICATION NUMBER: PCT/FR98/00870
;  PRIOR FILING DATE: 1998-04-29
;  NUMBER OF SEQ ID NOS: 34
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
;   LENGTH: 64
;   TYPE: PRT
;   ORGANISM: MSRV
US-09-403-343B-25

Alignment Scores:
Pred. No.:                0.127         Length:          64
Score:                    75.00         Matches:         16
Percent Similarity:       51.22%        Conservative:    5
Best Local Similarity:    39.02%        Mismatches:      16
Query Match:              23.15%        Indels:          4
DB:                       4             Gaps:            2

US-09-673-274B-2 (1-169) x US-09-403-343B-25 (1-64)

Qy          52 TGCAACGGTGAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTC------ 105
               |||:::|||   ||||||:::   |||   |||   |||   ||||||
Db          20 CysAspGlyLeuCysLysLysAsnGlyAlaLysSerGlySerCysGlyPheLeuValPro 39
```

```
)y         106 GCTAACGTGAACTGCTGGTGCGAGACTTGAGAGCTCGGCGAGGCGAACGTGTCGACGGAT 165
               :::   :::   |||||||||         :::|||   ||:               |||
b           40 SerGlyLeuAlaCysTrpCysAsn------AspLeuProGluAsnValProIleLysAsp 57

)y         166 CCG 168
               |||
)b          58 Pro 58

ESULT 9
JS-09-480-251-4
 Sequence 4, Application US/09480251
 Patent No. 6465719
 GENERAL INFORMATION:
  APPLICANT:   DeRose, Richard
  APPLICANT:  Freyssinet, Georges
  APPLICANT:  Hoffman, Jules
  TITLE OF INVENTION:  Chimeric Gene Encoding Drosomycin,
  TITLE OF INVENTION:  Vector Containing It And Production Of Disease-Resistant
  TITLE OF INVENTION:  Transgenic Plants
  FILE REFERENCE:   A32889-PCT-USA-A
  CURRENT APPLICATION NUMBER: US/09/480,251
  CURRENT FILING DATE:  2000-01-11
  PRIOR APPLICATION NUMBER:   PCT/FR98/01462
  PRIOR FILING DATE:   1998-07-08
  PRIOR APPLICATION NUMBER:   FRANCE 97/09,115
  PRIOR FILING DATE:   1997-07-11
  PRIOR APPLICATION NUMBER:   FRANCE 9709,663
  PRIOR FILING DATE:   1997-07-24
  NUMBER OF SEQ ID NOS:   15
  SOFTWARE:   FastSEQ for Windows Version 3.0
 SEQ ID NO 4
   LENGTH: 44
   TYPE: PRT
   ORGANISM: Drosophila melanogaster
JS-09-480-251-4

lignment Scores:
red. No.:               0.138          Length:          44
core:                   74.50          Matches:         15
ercent Similarity:      42.50%         Conservative:    2
est Local Similarity:   37.50%         Mismatches:      16
uery Match:             22.99%         Indels:          7
B:                      4              Gaps:            3

JS-09-673-274B-2 (1-169) x US-09-480-251-4 (1-44)

)y          13 GGTTCCTGC---GTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAG 69
               |||   |||   ||||||                        |||         ||||||
)b           9 GlyProCysAlaValTrpAsp------------AsnGluThrCysArgArgValCysLys 24

)y          70 AGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
                     |||         |||||||||            ::::::   ||||||||||||
)b          25 GluGluGlyArgSerSerGlyHisCysSer------ProSerLeuLysCysTrpCysGlu 42

ESULT 10
JS-09-480-251-6
 Sequence 6, Application US/09480251
 Patent No. 6465719
 GENERAL INFORMATION:
  APPLICANT:   DeRose, Richard
  APPLICANT:  Freyssinet, Georges
  APPLICANT:  Hoffman, Jules
  TITLE OF INVENTION:  Chimeric Gene Encoding Drosomycin,
  TITLE OF INVENTION:  Vector Containing It And Production Of Disease-Resistant
  TITLE OF INVENTION:  Transgenic Plants
  FILE REFERENCE:   A32889-PCT-USA-A
  CURRENT APPLICATION NUMBER: US/09/480,251
  CURRENT FILING DATE:  2000-01-11
  PRIOR APPLICATION NUMBER:   PCT/FR98/01462
  PRIOR FILING DATE:   1998-07-08
  PRIOR APPLICATION NUMBER:   FRANCE 97/09,115
  PRIOR FILING DATE:   1997-07-11
```

```
;  PRIOR APPLICATION NUMBER:   FRANCE 9709,663
;  PRIOR FILING DATE:   1997-07-24
;  NUMBER OF SEQ ID NOS:   15
;  SOFTWARE:   FastSEQ for Windows Version 3.0
; SEQ ID NO 6
;   LENGTH: 69
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: The ubiquitin - drosomycin fusion peptide
US-09-480-251-6

Alignment Scores:
Pred. No.:                0.149          Length:         69
Score:                    74.50          Matches:        15
Percent Similarity:       42.50%         Conservative:   2
Best Local Similarity:    37.50%         Mismatches:     16
Query Match:              22.99%         Indels:         7
DB:                       4              Gaps:           3

US-09-673-274B-2 (1-169) x US-09-480-251-6 (1-69)

Qy          13 GGTTCCTGC---GTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAG 69
               |||   |||   ||||||                        |||         ||||||
Db          34 GlyProCysAlaValTrpAsp------------AsnGluThrCysArgArgValCysLys 49

Qy          70 AGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
                     |||         |||||||||            ::::::   ||||||||||||
Db          50 GluGluGlyArgSerSerGlyHisCysSer------ProSerLeuLysCysTrpCysGlu 67

RESULT 11
US-09-480-251-2
; Sequence 2, Application US/09480251
; Patent No. 6465719
; GENERAL INFORMATION:
;  APPLICANT:   DeRose, Richard
;  APPLICANT:  Freyssinet, Georges
;  APPLICANT:  Hoffman, Jules
;  TITLE OF INVENTION:  Chimeric Gene Encoding Drosomycin,
;  TITLE OF INVENTION:  Vector Containing It And Production Of Disease-Resistant
;  TITLE OF INVENTION:  Transgenic Plants
;  FILE REFERENCE:   A32889-PCT-USA-A
;  CURRENT APPLICATION NUMBER: US/09/480,251
;  CURRENT FILING DATE:  2000-01-11
;  PRIOR APPLICATION NUMBER:   PCT/FR98/01462
;  PRIOR FILING DATE:   1998-07-08
;  PRIOR APPLICATION NUMBER:   FRANCE 97/09,115
;  PRIOR FILING DATE:   1997-07-11
;  PRIOR APPLICATION NUMBER:   FRANCE 9709,663
;  PRIOR FILING DATE:   1997-07-24
;  NUMBER OF SEQ ID NOS:   15
;  SOFTWARE:   FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 70
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-09-480-251-2

Alignment Scores:
Pred. No.:                0.149          Length:         70
Score:                    74.50          Matches:        15
Percent Similarity:       42.50%         Conservative:   2
Best Local Similarity:    37.50%         Mismatches:     16
Query Match:              22.99%         Indels:         7
DB:                       4              Gaps:           3

US-09-673-274B-2 (1-169) x US-09-480-251-2 (1-70)

Qy          13 GGTTCCTGC---GTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAG 69
               |||   |||   ||||||                        |||         ||||||
Db          35 GlyProCysAlaValTrpAsp------------AsnGluThrCysArgArgValCysLys 50

Qy          70 AGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
```

```
                    |:'         | !:||||!            ::::::    ||||'||||.|
Db          51 GluGluGlyArgSerSerGlyHisCysSer------ProSerLeuLysCysTrpCysGlu 68

RESULT 12
US-09-252-991A-26807
; Sequence 26807, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 26807
;   LENGTH: 158
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26807

Alignment Scores:
Pred. No.:               0.23         Length:         158
Score:                   73.50        Matches:        21
Percent Similarity:      47.46%       Conservative:   7
Best Local Similarity:   35.59%       Mismatches:     16
Query Match:             24.26%       Indels:         15
DB:                      4            Gaps:           4

US-09-673-274B-2 (1-169) x US-09-252-991A-26807 (1-158)

Qy         169 CCGGATCCGTCGACACGTTCGCCTCGCCGAGCTCTCAAGTCTCGCACCAGCAGTTCACGT 110
                ||   ':::::||!|.||   .|.|||            :;|   ::::::|||:::|||
Db          26 ProSerProAlaThrArgLysProArg------------SerProSerAlaSerAlaArg 41

Qy         109 TAGCGAAGGAACCGCAGTGACCACCCTTGTAACCCC---TCCTCTTGCACTCACCGTTGC 53
                           ::::::    ||||  :         :.|:::         |:,|||
Db          42 -- ---------LysThrSerSerProCysGlyThrGlySerAlaAlaAlaThrArgCys 57

Qy          52 AATCGG------------------AAGTGTAGTTCACAGCACCCCACACGCAGGAAC 14
                  '||                     ||     ||,      ||| ,'    : ||||:
Db          58 GlyArgProSerAlaCysArgSerProCysProSerAlaSerProThrSerArgAsn 76

RESULT 13
US-09-252-991A-22334
; Sequence 22334, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 22334
;   LENGTH: 145
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22334

Alignment Scores:
Pred. No.:               0.262        Length:         145
Score:                   73.00        Matches:        22
Percent Similarity:      45.76%       Conservative:   5
```

```
Best Local Similarity:   37.29%       Mismatches:     22
Query Match:             24.09%       Indels:         10
DB:                      4            Gaps:           3

US-09-673-274B-2 (1-169) x US-09-252-991A-22334 (1-145)

Qy         163 CCGTCGACACGTTCGCCTCGCCGAGCTCTCAAGTCT------CGCACCAGCAGTTCACGT 110
               :||||:   |||       |||      |||:::||      :::      :::   :::
Db          52 ProSerAlaArgLysAlaArgProArgLeuArgSerProProLysArgGlnAlaGlyLys 71

Qy         109 TAGCGAAGGAACCGCAGTGACCACCCTTGTAACCCCTCC---------------TCTTGC 65
                  ||||||               |||||||||||||||                ||||||
Db          72 ValArgArgAlaProArgArgArgProCysAsnProSerArgAlaProSerThrSerCys 91

Qy          64 ACTCACCGT---------TGCAATCGGAAGTGTAGTTCACAGCACCCCACACGCAGG 17
                     |||               |||:::|||||||||      |||      |||
Db          92 AlaCysArgArgSerProAlaArgArgArgCysSerSerProThrProProProArg 110

RESULT 14
US-09-252-991A-17266
; Sequence 17266, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 17266
;   LENGTH: 248
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17266

Alignment Scores:
Pred. No.:               0.445        Length:         248
Score:                   71.50        Matches:        18
Percent Similarity:      44.44%       Conservative:   6
Best Local Similarity:   33.33%       Mismatches:     17
Query Match:             23.60%       Indels:         13
DB:                      4            Gaps:           2

US-09-673-274B-2 (1-169) x US-09-252-991A-17266 (1-248)

Qy         163 CCGTCGACACGTTCGCCTCGCCGA---------GCTCTCAAGTCTCGCACCAGCAGTTCA 113
               |||:::   ||::::   ||||||            |||:::   |||         |||
Db          39 ProAlaArgArgAlaGluArgArgArgAlaAlaThrLeuArgGlnArgAspProGluSer 58

Qy         112 CGTTAGCGAAGGAACCGCAGTGACCACCCTTGTAACCCCTCCTCTTGCACTCACCGTTGC 53
               :.|      |:|   :||                                    ||||||
Db          59 ArgAlaAspArgArgArg------------------------------LeuArgArgCys 68

Qy          52 AATCGGAAGTGTAGTTCACAGCACCCCACACGCAGGAACCGA 11
                  |||:::|||:::  :::||||||   |||      |||
Db          69 ArgArgArgCysAlaGlyArgHisProAlaArgProGlyArg 82

RESULT 15
US-09-252-991A-26857
; Sequence 26857, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
```

```
  PRIOR APPLICATION NUMBER:  US 60/074,788
  PRIOR FILING DATE:  1998-02-18
  PRIOR APPLICATION NUMBER:  US 60/094,190
  PRIOR FILING DATE:  1998-07-27
  NUMBER OF SEQ ID NOS:  33142
 SEQ ID NO 26857
   LENGTH: 397
   TYPE: PRT
   ORGANISM: Pseudomonas aeruginosa
S-09-252-991A-26857

lignment Scores:
red. No.:               0.749          Length:        397
core:                   70.00          Matches:       20
ercent Similarity:      51.02%         Conservative:  5
est Local Similarity:   40.82%         Mismatches:    23
uery Match:             23.10%         Indels:        1
B:                      4              Gaps:          0

S-09-673-274B-2 (1-169) x US-09-252-991A-26857 (1-397)

y        169 CCGGATCCGTCGACACGTTCGCCTCGCCGAGCTCTCAAGTCTCGCACCAGCAGTTCACGT 110
             |:|      |||||||||   |||   ||:        |:.   ::::::|||
b        270 ProSerArgSerThrArgLysProSerArgLeuTyrCysSerAlaSerAlaSerArgAla 289

y        109 TAGCGAAGGAACCGCAGTGACCACCCTTGTAACCCCTCCTCTTGCACTCACCGTTGCAAT 50
                ||||||:::  |||               ||||||   :||||:    :||||:||
b        290 ArgArgArgSerSerSerSerGlyArgSerThrProSerTrpCysThrArg-ValAlaIl 309

y         49 CGGAAGTGTAGTTCACAGCACCCCA 25
             |   |||:::     :::   |||
b        309 eSerSerIleSerSerThrAlaPro 317

earch completed: October 16, 2003, 17:13:05
ob time : 18.8439 secs
```

```
                    GenCore version 5.1.6
             Copyright (c) 1993 - 2003  Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on:         October 16, 2003, 17:06:37 ; Search time 16.2259 Seconds
                                             (without alignments)
                                             1564.690 Million cell updates/sec

Title:          US-09-673-274B-2_COPY_1_132
Perfect score:  267
Sequence:       1 gataagcttatcggttcctg..........tgaactgctggtgcgagact 132

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters:      566616

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09673274/runat_16102003_170628_2555/app_query.fasta_1.654
-DB=PIR_76 -QFMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09673274_@CGN_1_1_77_@runat_16102003_170628_2555 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      PIR_76:*
                1:  pir1:*
                2:  pir2:*
                3:  pir3:*
                4:  pir4:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
                   %
Result             Query
  No.     Score    Match  Length DB  ID                 Description
-------------------------------------------------------------------------------
     1       86     32.2     66  2   A29386             neurotoxin III - s
     2     80.5     30.1     65  2   A59222             neurotoxin X - sco
     3     79.5     29.8     66  2   A55869             crustacean-specifi
     4       79     29.6     85  2   S68906             alpha-toxin Bot XI
     5     78.5     29.4     84  1   NTSR3A             neurotoxin III pre
     6     77.5     29.0     65  1   NTSR4L             neurotoxin IV - Eg
     7     76.5     28.7     61  2   B34123             depressant insect
     8       76     28.5     66  1   NTSR4E             neurotoxin M14 - l
     9     75.5     28.3     65  1   NTSR1C             neurotoxin 1 - bar
    10     74.5     27.9     70  2   A55824             drosomycin precurs
    11       74     27.7     59  2   C23727             neurotoxin V-5 - b
    12       74     27.7     64  2   JC1321             neurotoxin IV - Sa
    13       73     27.3     64  1   NTSR5L             neurotoxin V - Egy
    14     70.5     26.4     65  1   NTSR3C             neurotoxin 3 - bar
```

```
    15     70.5     26.4     87  2   JN0670             Na+-channel-blocki
    16     69.5     26.0     84  1   NTSR2N             Na+-channel-blocki
    17     68.5     25.7     62  2   D59352             depressant insect-
    18     68.5     25.7     66  1   NTSR9E             neurotoxin M9 - le
    19     68.5     25.7     83  1   NTSRIA             neurotoxin I precu
    20       68     25.5     64  1   NTSR5M             neurotoxin V - sco
    21       68     25.5     77  2   T14395             proteinase inhibit
    22     67.5     25.3     61  2   B59352             depressant insect-
    23     67.5     25.3     61  2   C59352             depressant insect-
    24     67.5     25.3     83  1   NTSRI2             neurotoxin I' prec
    25     67.5     25.3     87  2   JN0669             Na+-channel-blocki
    26       67     25.1     86  2   JN0671             Na+-channel-blocki
    27     66.5     24.9     65  1   NTSR1B             neurotoxin I - sco
    28     66.5     24.9     66  1   NTSR2C             neurotoxin 2 - bar
    29       66     24.7   3672  2   T23433             hypothetical prote
    30       66     24.7   3704  2   T37316             probable laminin a
    31     65.5     24.5     61  2   A59006             excitatory insect
    32     64.5     24.2     66  2   B23727             neurotoxin V-4 - b
    33     64.5     24.2     85  2   A61616             depressant insect
    34       64     24.0     77  2   S30578             protease inhibitor
    35       64     24.0     78  2   S52634             gamma-thionin - Pe
    36       64     24.0     92  2   S43228             defensin - fruit f
    37     63.5     23.8     85  2   A40472             depressant insect
    38       63     23.6     64  2   A35940             neurotoxin alpha-I
    39       63     23.6     85  2   A39306             alpha insect toxin
c   40       63     26.8   4753  1   A47437             LDL-receptor-relat
    41       62     23.2   3712  2   S18253             laminin alpha-1 ch
    42     61.5     23.0     65  1   NTSR2B             neurotoxin II - sc
    43     61.5     23.0     65  2   A38394             neurotoxin AaH IT4
    44     61.5     23.0     66  2   A27406             toxin CssII - Mexi
    45     61.5     23.0     84  2   JE0143             anti-mammals neuro

                              ALIGNMENTS

RESULT 1
A29386
neurotoxin III - scorpion (Buthus occitanus)
C;Species: Buthus occitanus mardochei
C;Date: 21-May-1988 #sequence_revision 21-May-1988 #text_change 07-Feb-1997
C;Accession: A29386
R;Vargas, O.; Martin, M.F.; Rochat, H.
Eur. J. Biochem. 162, 589-599, 1987
A;Title: Characterization of six toxins from the venom of the Moroccan scorpion Buthus
A;Reference number: A29386; MUID:87161829; PMID:3104036
A;Accession: A29386
A;Molecule type: protein
A;Residues: 1-66 <VAR>
C;Superfamily: scorpion neurotoxin
C;Keywords: neurotoxin

Alignment Scores:
Pred. No.:              0.0247        Length:        66
Score:                  86.00         Matches:       15
Percent Similarity:     55.00%        Conservative:  7
Best Local Similarity:  37.50%        Mismatches:    16
Query Match:            32.21%        Indels:        2
DB:                     2             Gaps:          1

US-09 673-274B-2_COPY_1_132 (1-132) x A29386 (1-66)

Qy         16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
                 :::||||||:::                  :::||'  ||':::      ||'|||    :::
Db         11 AsnCysValTyrHisCysPheProGlySerSerGlyCysAspThrLeuCysLysGluLys 30

Qy         76 GGTTACAAGGGTGGTCACTGCGGTTCCTTC------GCTAACGTGAACTGCTGGTGCGAG 129
                 |||        |||||:|||:|||               :::   |||   ||:|'|||| :::
Db         31 GlyAlaThrSerGlyHisCysGlyPheLeuProGlySerGlyValAlaCysTrpCysAsp 50

RESULT 2
A59222
neurotoxin X - scorpion (Leiurus quinquestriatus)
C;Species: Leiurus quinquestriatus hebraeus
```

```
C;Date: 31 Dec-2001 #sequence_revision 31-Dec-2001 #text_change 19-Apr-2002
C;Accession: A59222
R;Corzo, G.; Escoubas, P.; Nakajima, T.
submitted to the Protein Sequence Database, April 2000
A;Description: A procedure for the rapid purification of non commercial and new scorpion
A;Reference number: A59222
A;Accession: A59222
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-65 <COR>
C;Superfamily: scorpion neurotoxin
C;Keywords: amidated carboxyl end; neurotoxin; venom
F;13-64,17-37,23-47,27-49/Disulfide bonds: #status predicted
F;65/Modified site: amidated carboxyl end (Arg) #status predicted

Alignment Scores:
Pred. No.:              0.117        Length:        65
Score:                  80.50        Matches:       17
Percent Similarity:     57.14%       Conservative:  7
Best Local Similarity:  40.48%       Mismatches:    9
Query Match:            30.15%       Indels:        9
DB:                     2            Gaps:          4

US-09 673-274B-2_COPY_1_132 (1-132) x A59222 (1-65)

Qy          16 TCCTGCGTGTGG------GGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAG 69
               :::|'||||:::      |,||||   :::.||         |||::    ! |||:
Db          12 AsnCysValTyrThrCysGlyAlaAsnSerTyr---------CysAsnThrGluCysThr 28

Qy          70 AGGAGGGGTTACAAGGGTGGTCACTGC---------GGTTCCTTCGCTAACGTGAACTGC 120
               :::   |||   :::   |'|:::  |          ||   :::   |||      |||
Db          29 LysAsnGlyAlaGluSerGlyTyrCysGlnTrpPheGlyLysTyrGlyAsnAla---Cys 47

Qy         121 TGGTGC 126
               || : |
Db          48 TrpCys 49

RESULT 3
A55869
crustacean specific toxin 1 - scorpion (Centruroides limpidus)
C;Species: Centruroides limpidus limpidus
C;Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 17-Mar-1999
C;Accession: A55869
R;Lebreton, F.; Delepierre, M.; Ramirez, A.N.; Balderas, C.; Possani, L.D.
Biochemistry 33, 11135-11149, 1994
A;Title: Primary and NMR three-dimensional structure determination of a novel crustacean
A;Reference number: A55869; MUID:95244424; PMID:7727365
A;Accession: A55869
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-66 <LEB>
C;Superfamily: scorpion neurotoxin
C;Keywords: disulfide bond; toxin
F;12 65,16-41,25-46,29-48/Disulfide bonds: #status experimental

Alignment Scores:
Pred. No.:              0.155        Length:        66
Score:                  79.50        Matches:       17
Percent Similarity:     53.85%       Conservative:  4
Best Local Similarity:  43.59%       Mismatches:    11
Query Match:            29.78%       Indels:        7
DB:                     2            Gaps:          3

US 09 673-274B-2_COPY_1_132 (1-132) x A55869 (1-66)

Qy          19 TGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGGGGT 78
               |.|   |||         |||         :::|||:::   |||||:|||   :::
Db          16 CysPheTrpLeuGlyLysAsn------GluAsnCysAspLysGluCysLysAlaLysAsn 33

Qy          79 TACAAGGGT------GGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
                     |||      |||:::|||   ||||:.|||         |. |||:|||||
Db          34 GlnGlyGlySerTyrGlyTyrCysTyrSerPheAla---------CysTrpCysGlu 49
```

```
RESULT 4
S68906
alpha-toxin Bot XIV precursor - scorpion (Buthus occitanus)
C;Species: Buthus occitanus tunetanus
C;Date: 23-Jul-1997 #sequence_revision 01-Aug-1997 #text_change 16-Jul-1999
C;Accession: S68906
R;Bouhaouala-Zahar, B.; Ducancel, F.; Zenouaki, I.; ben Khalifa, R.; Borchani, L.; Pel
Eur. J. Biochem. 238, 653-660, 1996
A;Title: A recombinant insect-specific alpha-toxin of Buthus occitanus tunetanus scorp
A;Reference number: S68906; MUID:96300228; PMID:8706664
A;Accession: S68906
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-85 <BOU>
A;Cross-references: GB:X92376; NID:g1041277; PIDN:CAA63120.1; PID:g1041278
C;Superfamily: scorpion neurotoxin
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-85/Product: alpha-toxin #status predicted <MAT>

Alignment Scores:
Pred. No.:              0.171        Length:        85
Score:                  79.00        Matches:       14
Percent Similarity:     50.00%       Conservative:  6
Best Local Similarity:  35.00%       Mismatches:    18
Query Match:            29.59%       Indels:        2
DB:                     2            Gaps:          1

US-09-673-274B-2_COPY_1_132 (1-132) x S68906 (1-85)

Qy          16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
               :::||'   :::      :::      :::|||   |||:::      ||||||
Db          29 AsnCysAlaTyrHisCysLeuLysIleSerSerGlyCysAspThrLeuCysLysGluAsn 48

Qy          76 GGTTACAAGGGTGGTCACTGCGGT------TCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
               |||         ||||.||||:||      ||:               |||||||||:::
Db          49 GlyAlaThrSerGlyHisCysGlyHisLysSerGlyHisGlySerAlaCysTrpCysLys 68

RESULT 5
NTSR3A
neurotoxin III precursor - Sahara scorpion
N;Alternate names: mammalian neurotoxin II
C;Species: Androctonus australis (Sahara scorpion)
C;Date: 30-Sep-1979 #sequence_revision 04-Oct-1996 #text_change 18-Jun-1999
C;Accession: C34444; A01740
R;Bougis, P.E.; Rochat, H.; Smith, L.A.
J. Biol. Chem. 264, 19259-19265, 1989
A;Title: Precursors of Androctonus australis scorpion neurotoxins. Structures of precu
A;Reference number: A34444; MUID:90037062; PMID:2808423
A;Accession: C34444
A;Molecule type: mRNA
A;Residues: 1-84 <BOU>
A;Cross-references: GB:M27703; GB:J05102; NID:g161142; PIDN:AAA29948.1; PID:g161143
A;Experimental source: clone pcD-634
R;Kopeyan, C.; Martinez, G.; Rochat, H.
Eur. J. Biochem. 94, 609-615, 1979
A;Title: Amino acid sequence of neurotoxin III of the scorpion Androctonus australis H
A;Reference number: A01740; MUID:79148414; PMID:428402
A;Accession: A01740
A;Molecule type: protein
A;Residues: 20-26,'N',28-83 <KOP>
C;Superfamily: scorpion neurotoxin
C;Keywords: neurotoxin; venom
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-83/Product: neurotoxin III #status experimental <MAT>
F;31-81,35-53,39-63,43-65/Disulfide bonds: #status predicted

Alignment Scores:
Pred. No.:              0.197        Length:        84
Score:                  78.50        Matches:       15
Percent Similarity:     53.85%       Conservative:  6
Best Local Similarity:  38.46%       Mismatches:    13
Query Match:            29.40%       Indels:        5
DB:                     1            Gaps:          2
```

```
S-09-673-274B-2_COPY_1_132 (1-132) x NTSR3A (1-84)

y          16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
              :::||||||:::      |||               |:|:::|||   |:||||:::
b          30 AsnCysValTyrHisCysVal---------ProProCysAspGlyLeuCysLysLysAsn 46

)y         76 GGTTACAAGGGTGGTCACTGCGGTTCCTTC------GCTAACGTGAACTGCTGGTGC 126
              |||   |||   |||   ||||||            ::: \ :::   |:|||||||
)b         47 GlyAlaLysSerGlySerCysGlyPheLeuIleProSerGlyLeuAlaCysTrpCys 65

 ESULT 6
ITSR4L
ieurotoxin IV - Egyptian scorpion
:;Species: Leiurus quinquestriatus (Egyptian scorpion)
:;Date: 15-Oct-1982 #sequence_revision 27-Nov-1985 #text_change 28-Aug-1998
:;Accession: A01749
l;Kopeyan, C.; Martinez, G.; Rochat, H.
'EBS Lett. 181, 211-217, 1985
i;Title: Primary structure of toxin IV of Leiurus quinquestriatus quinquestriatus.
i;Reference number: A01749
i;Accession: A01749
i;Molecule type: protein
i;Residues: 1-65 <KOP>
i;Note: the source is designated as Leiurus quinquestriatus quinquestriatus
:;Superfamily: scorpion neurotoxin
:;Keywords: blocked carboxyl end; neurotoxin; sodium channel inhibitor; venom
';13-64,17-37,23-47,27-49/Disulfide bonds: #status predicted
';65/Modified site: blocked carboxyl end (Arg) (probably amidated) #status experimental

ilignment Scores:
'red. No.:                0.274          Length:          65
 core:                    77.50          Matches:         16
'ercent Similarity:       57.14%         Conservative:    8
lest Local Similarity:    38.10%         Mismatches:      9
)uery Match:              29.03%         Indels:          9
)B:                       1              Gaps:            4

)S-09-673-274B-2_COPY_1_132 (1-132) x NTSR4L (1-65)

)y         16 TCCTGCGTGTGG------GGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAG 69
              :::||||||:::      |||:::   :::'||         ||||:|    .||||;
)b         12 AsnCysValTyrThrCysGlySerAsnSerTyr---------CysAsnThrGluCysThr 28

)y         70 AGGAGGGGTTACAAGGGTGGTCACTGC---------GGTTCCTTCGCTAACGTGAACTGC 120
              :::   |||   :::   |||:::|||         |||   :::   ||'      |||
)b         29 LysAsnGlyAlaGluSerGlyTyrCysGlnTrpLeuGlyLysTyrGlyAsnAla---Cys 47

)y        121 TGGTGC 126
              ||||||
)b         48 TrpCys 49

lESULT 7
334123
lepressant insect toxin LqqIT2 - scorpion (Leiurus quinquestriatus)
:;Species: Leiurus quinquestriatus quinquestriatus
:;Date: 06-Jul-1990 #sequence_revision 06-Jul-1990 #text_change 02-Sep-2000
:;Accession: B34123; S08268; B38526
l;Zlotkin, E.; Fowler, E.; Eitan, M.; Moyer, M.; Adams, M.E.
Toxicon 28, 170, 1990
\;Title: On the chemistry and action of the depressant insect toxins.
\;Reference number: A34123
\;Accession: B34123
\;Molecule type: protein
\;Residues: 1-61 <ZLO1>
l;Kopeyan, C.; Mansuelle, P.; Sampieri, F.; Brando, T.; Bahraoui, E.M.; Rochat, H.; Grar
'EBS Lett. 261, 423-426, 1990
\;Title: Primary structure of scorpion anti-insect toxins isolated from the venom of Lei
\;Reference number: S08267; MUID:90184494; PMID:2311768
\;Accession: S08268
\;Molecule type: protein
\;Residues: 1-61 <KOP>
l;Zlotkin, E.; Eitan, M.; Bindokas, V.P.; Adams, M.E.; Moyer, M.; Burkhart, W.; Fowler,
```

```
Biochemistry 30, 4814-4821, 1991
A;Title: Functional duality and structural uniqueness of depressant insect selective ne
A;Reference number: A38526; MUID:91230120; PMID:2029523
A;Accession: B38526
A;Molecule type: protein
A;Residues: 1-61 <ZLO2>
A;Cross-references: CAS:130300-67-3
C;Superfamily: scorpion neurotoxin
C;Keywords: neurotoxin; venom
F;10-60,14-35,21-42,25-44/Disulfide bonds: #status predicted

Alignment Scores:
Pred. No.:                0.368          Length:          61
Score:                    76.50          Matches:         15
Percent Similarity:       55.26%         Conservative:    6
Best Local Similarity:    39.47%         Mismatches:      12
Query Match:              28.65%         Indels:          5
DB:                       2              Gaps:            2

US-09-673-274B-2_COPY_1_132 (1-132) x B34123 (1-61)

Qy         16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
              |||||  ::::::|||                    .|||||   |||:'||||
Db         13 SerCysLeuPheGly------------AsnGluGlyCysAsnLysGluCysLysSerTyr 28

Qy         76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
              |||         |||:::|||   ::::::      :::   ||||||||:|||
Db         29 GlyGlySerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 45

RESULT 8
NTSR4E
neurotoxin M14 - lesser Asian scorpion
C;Species: Mesobuthus eupeus (lesser Asian scorpion)
C;Date: 30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change 31-Dec-1993
C;Accession: JT0020
R;Volkova, T.M.; Garsia, A.F.; Telezhinskaya, I.N.; Potapenko, N.A.; Grishin, E.V.
Bioorg. Khim. 11, 1445-1456, 1985
A;Title: Study of neurotoxins from the venom of central Asian scorpion Buthus eupeus.
A;Reference number: A94652; MUID:86130759; PMID:4091860
A;Accession: JT0020
A;Molecule type: protein
A;Residues: 1-66 <VOL>
C;Comment: The venom of this scorpion contains three major polypeptide neurotoxins: M9,
C;Superfamily: scorpion neurotoxin
C;Keywords: neurotoxin; venom
F;12-65,16-36,22-46,26-48/Disulfide bonds: #status predicted

Alignment Scores:
Pred. No.:                0.418          Length:          66
Score:                    76.00          Matches:         13
Percent Similarity:       58.62%         Conservative:    4
Best Local Similarity:    44.83%         Mismatches:      8
Query Match:              28.46%         Indels:          4
DB:                       1              Gaps:            2

US-09-673-274B-2_COPY_1_132 (1-132) x NTSR4E (1-66)

Qy         52 TGCAACGGTGAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGC---------GGTTCC 102
              '||:::   :|||||| |:::   ||         :||   :::|||         |||
Db         22 CysAspSerGluCysLysLysAsnGlyAlaAspGlySerTyrCysGlnTrpLeuGlyArg 41

Qy        103 TTCGCTAACGTGAACTGCTGGTGCGAG 129
              |||   |||      |||:||||||:::
Db         42 PheGlyAsnAla---CysTrpCysLys 49

RESULT 9
NTSR1C
neurotoxin 1 - bark scorpion (tentative sequence)
C;Species: Centruroides sculpturatus (bark scorpion)
C;Date: 22-Jun-1981 #sequence_revision 22-Jun-1981 #text_change 31 Mar 2000
C;Accession: A01752
R;Babin, D.R.; Watt, D.D.; Goos, S.M.; Mlejnek, R.V.
Arch. Biochem. Biophys. 164, 694-706, 1974
```

```
A;Title: Amino acid sequences of neurotoxic protein variants from the venom of Centruroi
A;Reference number: A90058; MUID:75163395; PMID:4460885
A;Accession: A01752
A;Molecule type: protein
A;Residues: 1-65 <BAB>
C;Superfamily: scorpion neurotoxin
C;Keywords: neurotoxin; venom
F;12-64,16-41,27-46,29-48/Disulfide bonds: #status predicted

Alignment Scores:
Pred. No.:               0.483        Length:          65
Score:                   75.50        Matches:         14
Percent Similarity:      51.35%       Conservative:    5
Best Local Similarity:   37.84%       Mismatches:      15
Query Match:             28.28%       Indels:          3
DB:                      1            Gaps:            1

US-09-673-274B-2_COPY_1_132 (1-132) x NTSR1C (1-65)

Qy          19 TGCCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGGGGT 78
               |||   |||         ||:      :::   | |       :::   |||   :::|:|
Db          16 CysPheTrpLeuGlyLysAsnGluHisAsnThrCysGluCysLysAlaLysAsnGlnGly 35

Qy          79 TACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
                        |||:::|||   :::|||:.:         :.|||.:|||.
Db          36 GlySerTyrGlyTyrCysTyrAlaPheAla---------CysTrpCysGlu 49

RESULT 10
A55824
drosomycin precursor - fruit fly (Drosophila melanogaster)
N;Alternate names: cysteine-rich peptide
C;Species: Drosophila melanogaster
C;Date: 23-Mar-1995 #sequence_revision 05-Apr-1995 #text_change 21-Jul-2000
C;Accession: A55824; S38575
R;Fehlbaum, P.; Bulet, P.; Michaut, L.; Lagueux, M.; Broekaert, W.F.; Hetru, C.; Hoffman
J. Biol. Chem. 269, 33159-33163, 1994
A;Title: Insect immunity. Septic injury of Drosophila induces the synthesis of a potent
A;Reference number: A55824; MUID:95105209; PMID:7806546
A;Accession: A55824
A;Molecule type: mRNA
A;Residues: 1-70 <FEH>
A;Cross-references: EMBL:X75595; NID:g414661; PIDN:CAA53267.1; PID:g414662
A;Note: the mature peptide sequence was confirmed by peptide sequencing
C;Comment: The exact site of signal peptide cleavage (and the requirement, if any, for a
C;Genetics:
A;Gene: FlyBase:Drs
A;Cross-references: FlyBase:FBgn0010381
C;Keywords: disulfide bond; extracellular protein
F;27-70/Product: drosomycin #status experimental <MAT>

Alignment Scores:
Pred. No.:               0.632        Length:          70
Score:                   74.50        Matches:         15
Percent Similarity:      42.50%       Conservative:    2
Best Local Similarity:   37.50%       Mismatches:      16
Query Match:             27.90%       Indels:          7
DB:                      2            Gaps:            3

US-09-673-274B-2_COPY_1_132 (1-132) x A55824 (1-70)

Qy          13 GGTTCCTGC---GTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAG 69
               |||   |||   ||||||                        |::         ||||||
Db          35 GlyProCysAlaValTrpAsp------------AsnGluThrCysArgArgValCysLys 50

Qy          70 AGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
                     |||         |||||.|||            ::::::   ||||||||||||
Db          51 GluGluGlyArgSerSerGlyHisCysSer------ProSerLeuLysCysTrpCysGlu 68

RESULT 11
C23727
neurotoxin V-5 - bark scorpion
C;Species: Centruroides sculpturatus (bark scorpion)
C;Date: 31-Jan-1992 #sequence_revision 31-Jan-1992 #text_change 07-Feb-1997
C;Accession: C23727
R;David, R.M.; Krishna, N.R.; Watt, D.D.
Toxicon 29, 645-662, 1991
A;Title: Characterization of cationic binding sites of neurotoxins from venom of the
A;Reference number: A23727; MUID:92023284; PMID:1926166
A;Accession: C23727
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-59 <DAV>
C;Superfamily: scorpion neurotoxin
C;Keywords: neurotoxin

Alignment Scores:
Pred. No.:               0.751        Length:          59
Score:                   74.00        Matches:         16
Percent Similarity:      60.53%       Conservative:    7
Best Local Similarity:   42.11%       Mismatches:      7
Query Match:             27.72%       Indels:          8
DB:                      2            Gaps:            3

US-09-673-274B-2_COPY_1_132 (1-132) x C23727 (1-59)

Qy          16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
               |:|||||||      |||   ||||||         |||:::   :::||||||   :::
Db          14 SerCysVal------AlaAsnAsnTyr---------CysAspAsnGlnCysLysMetLys 28

Qy          76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
                        ||.|||:.||||   :::         ::::::|||:::||||||
Db          29 LysAlaSerGlyGlyHisCysTyrAla---------MetSerCysTyrCysGlu 43

RESULT 12
JC1321
neurotoxin IV - Sahara scorpion
C;Species: Androctonus australis (Sahara scorpion)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C;Accession: JC1321
R;Mansuelle, P.; Martin, M.F.; Rochat, H.; Granier, C.
Natural Toxins 1, 61-69, 1992
A;Title: The amino acid sequence of toxin IV from the Androctonus australis scorpion:
A;Reference number: JC1321; MUID:94221398; PMID:1344902
A;Accession: JC1321
A;Molecule type: protein
A;Residues: 1-64 <MAN>
C;Superfamily: scorpion neurotoxin
C;Keywords: neurotoxin

Alignment Scores:
Pred. No.:               0.74         Length:          64
Score:                   74.00        Matches:         12
Percent Similarity:      59.26%       Conservative:    4
Best Local Similarity:   44.44%       Mismatches:      9
Query Match:             27.72%       Indels:          2
DB:                      2            Gaps:            1

US-09-673-274B-2_COPY_1_132 (1-132) x JC1321 (1-64)

Qy          52 TGCAACGGTGAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTC------ 105
               :||:::|:|   :||||::::   |||   |||   |||   ||||||
Db          20 CysAspGlyLeuCysLysLysAsnGlyAlaLysSerGlySerCysGlyPheLeuValPro 39

Qy         106 GCTAACGTGAACTGCTGGTGC 126
               :::   :::   |||||||||
Db          40 SerGlyLeuAlaCysTrpCys 46

RESULT 13
NTSR5L
neurotoxin V - Egyptian scorpion
C;Species: Leiurus quinquestriatus (Egyptian scorpion)
C;Date: 30-Apr-1979 #sequence_revision 13-Jul-1981 #text_change 15-Oct-1996
C;Accession: A01741
R;Kopeyan, C.; Martinez, G.; Rochat, H.
FEBS Lett. 89, 54-58, 1978
A;Title: Amino acid sequence of neurotoxin V from the scorpion Leiurus quinquestriatu
```

```
A;Reference number: A01741; MUID:78191225; PMID:658402
A;Note: L. q. quinquestriatus
A;Accession: A01741
A;Molecule type: protein
A;Residues: 1-64 <KOP>
C;Superfamily: scorpion neurotoxin
C;Keywords: blocked carboxyl end; neurotoxin; venom
F;12-63,16-36,22-46,26-48/Disulfide bonds: #status predicted
F;64/Modified site: blocked carboxyl end (Asn) (probably amidated) #status experimental

Alignment Scores:
Pred. No.:              0.981         Length:        64
Score:                  73.00         Matches:       12
Percent Similarity:     60.71%        Conservative:  5
Best Local Similarity:  42.86%        Mismatches:    7
Query Match:            27.34%        Indels:        4
DB:                     1             Gaps:          2

US-09-673-274B-2_COPY_1_132 (1-132) x NTSR5L (1-64)

Qy         52 TGCAACGGTGAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGC---------GGTTCC 102
              ||||||    |||||||||::::::|:|    :::    |||:::|||
Db         22 CysAsnAspGluCysLysLysLysGlyGlyGluSerGlyTyrCysGlnTrpAlaSerPro 41

Qy        103 TTCGCTAACGTGAACTGCTGGTGC 126
              :::   |||        |||||||||
Db         42 TyrGlyAsnAla---CysTrpCys 48

RESULT 14
NTSR3C
neurotoxin 3 - bark scorpion
C;Species: Centruroides sculpturatus (bark scorpion)
C;Date: 15-Oct-1982 #sequence_revision 15-Oct-1982 #text_change 23-Aug-1996
C;Accession: A90058; A94470; A01754
R;Babin, D.R.; Watt, D.D.; Goos, S.M.; Mlejnek, R.V.
Arch. Biochem. Biophys. 164, 694-706, 1974
A;Title: Amino acid sequences of neurotoxic protein variants from the venom of Centruroi
A;Reference number: A90058; MUID:75163395; PMID:4460885
A;Accession: A90058
A;Molecule type: protein
A;Residues: 1-24,'NTC',28-63,'CS' <BAB>
R;Bhown, A.; Mole, J.
unpublished results, cited by Fontecilla-Camps, J.C., et al., Toxicon 20, 1-7, 1982
A;Reference number: A94470
A;Accession: A94470
A;Molecule type: protein
A;Residues: 1-65 <BHO>
R;Fontecilla-Camps, J.C.; Almassy, R.J.; Suddath, F.L.; Bugg, C.E.
Toxicon 20, 1-7, 1982
A;Title: The three-dimensional structure of scorpion neurotoxins.
A;Reference number: A94314; MUID:82200153; PMID:7080025
A;Contents: annotation; X-ray crystallography, 1.8 angstroms; disulfide bonds
A;Note: X-ray crystallographic studies were based on the revised sequence shown
C;Superfamily: scorpion neurotoxin
C;Keywords: neurotoxin; venom
F;12-65,16-41,25-46,29-48/Disulfide bonds: #status experimental

Alignment Scores:
Pred. No.:              1.98          Length:        65
Score:                  70.50         Matches:       15
Percent Similarity:     51.22%        Conservative:  6
Best Local Similarity:  36.59%        Mismatches:    13
Query Match:            26.40%        Indels:        7
DB:                     1             Gaps:          3

US-09-673-274B-2_COPY_1_132 (1-132) x NTSR3C (1-65)

Qy         19 TGCGTGTGGGGTGCTGTGAACTAC------ACTTCCGATTGCAACGGTGAGTGCAAGAGG 72
              |||   :::|||   :::                     |||:::   |||||||||
Db         12 CysLysTyrGlyCysLeuLysLeuGlyGluAsnGluGlyCysAspThrGluCysLysAla 31

Qy         73 AGGGGTTACAAGGGT------GGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGC 126
              :::         |||      |||:::|||   :::||||||         |||||||||
```

```
Db         32 LysAsnGlnGlyGlySerTyrGlyTyrCysTyrAlaPheAla---------CysTrpCys 48

Qy        127 GAG 129
              |:|
Db         49 Glu 49

RESULT 15
JN0670
Na+-channel-blocking toxin (clone cngtIII) precursor - scorpion (Centruroides noxius)
C;Species: Centruroides noxius
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 16-Jul-1999
C;Accession: JN0670
R;Becerril, B.; Vazquez, A.; Garcia, C.; Corona, M.; Bolivar, F.; Possani, L.D.
Gene 128, 165-171, 1993
A;Title: Cloning and characterization of cDNAs that code for Na+-channel blocking toxin
A;Reference number: JN0669; MUID:93292983; PMID:8390386
A;Accession: JN0670
A;Molecule type: mRNA
A;Residues: 1-87 <BEC>
A;Cross-references: GB:L05061; NID:g304568; PIDN:AAA28286.1; PID:g304569
A;Experimental source: venom gland
C;Superfamily: scorpion neurotoxin
C;Keywords: toxin
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-87/Product: Na+-channel-blocking toxin (clone cngtIII) #status predicted <MAT>

Alignment Scores:
Pred. No.:              1.88          Length:        87
Score:                  70.50         Matches:       15
Percent Similarity:     48.72%        Conservative:  4
Best Local Similarity:  38.46%        Mismatches:    13
Query Match:            26.40%        Indels:        7
DB:                     2             Gaps:          3

US-09-673-274B-2_COPY_1_132 (1-132) x JN0670 (1-87)

Qy         19 TGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGGGGT 78
              |:|   :||         :||                ||:::   ||:||||::    :::
Db         35 CysPheTrpLeuGlyLysAsn------GluGlyCysAspLysGluCysLysAlaLysAsn 52

Qy         79 TACAAGGGT------GGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
                    ||      |:::::|:    :::||              ||||::|||||:
Db         53 GlnGlyGlySerTyrGlyTyrCysTyrAlaPhe --------GlyCysTrpCysGlu 68


Search completed: October 16, 2003, 17:11:59
Job time : 17.2259 secs
```

```
                 GenCore version 5.1.6
          Copyright (c) 1993 - 2003  Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on:          October 16, 2003, 17:06:37 ; Search time 39.0299 Seconds
                                              (without alignments)
                                              1745.479 Million cell updates/sec

Title:           US-09-673-274B-2_COPY_1_132
Perfect score:   267
 equence:        1 gataagcttatcggttcctg..........tgaactgctggtgcgagact 132

Scoring table:   BLOSUM62
                 Xgapop 10.0 , Xgapext  0.5
                 Ygapop 10.0 , Ygapext  0.5
                 Fgapop  6.0 , Fgapext  7.0
                 Delop   6.0 , Delext   7.0

Searched:        830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters:      1661050

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09673274/runat_16102003_170627_2545/app_query.fasta_1.654
-DB=SPTREMBL_23 -QFMT=fastan -SUFFIX=rspt -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09673274_@CGN_1_1_216_@runat_16102003_170627_2545 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :       SPTREMBL_23:*
                 1:  sp_archea:*
                 2:  sp_bacteria:*
                 3:  sp_fungi:*
                 4:  sp_human:*
                 5:  sp_invertebrate:*
                 6:  sp_mammal:*
                 7:  sp_mhc:*
                 8:  sp_organelle:*
                 9:  sp_phage:*
                 10:  sp_plant:*
                 11:  sp_rodent:*
                 12:  sp_virus:*
                 13:  sp_vertebrate:*
                 14:  sp_unclassified:*
                 15:  sp_rvirus:*
                 16:  sp_bacteriap:*
                 17:  sp_archeap:*

    Pred. No. is the number of results predicted by chance to have a
    score greater than or equal to the score of the result being printed,
    and is derived by analysis of the total score distribution.

                                  SUMMARIES

Result            Query
  No.   Score   Match Length DB  ID                      Description
--------------------------------------------------------------------------
    1    84.5    31.6     65  5  Q17231                  Q17231 mesobuthus
    2    83.5    31.3     65  5  Q17230                  Q17230 mesobuthus
    3    83.5    31.3     85  5  Q8I0K7                  Q8i0k7 mesobuthus
    4    81.5    30.5     85  5  Q95WX6                  Q95wx6 mesobuthus
    5    81.5    30.5     85  5  Q9XY87                  Q9xy87 mesobuthus
    6    80.5    30.1     85  5  Q8T3T0                  Q8t3t0 mesobuthus
    7      80    30.0     91  10  P82761                  P82761 arabidopsis
    8    78.5    29.4     69  5  Q9VZQ5                  Q9vzq5 drosophila
    9    76.5    28.7     82  5  Q8MVS7                  Q8mvs7 leiurus qui
   10    76.5    28.7     85  5  Q9Y1U3                  Q9y1u3 mesobuthus
   11    74.5    27.9     69  5  Q9VZR2                  Q9vzr2 drosophila
   12    74.5    27.9     84  5  Q8WRY4                  Q8wry4 centruroide
   13      74    27.7     85  5  Q9GUA7                  Q9gua7 mesobuthus
   14    73.5    27.5    106  5  Q9VZR3                  Q9vzr3 drosophila
   15      72    27.0     85  5  Q95P69                  Q95p69 mesobuthus
   16    71.5    26.8     73  10  Q9FPG1                  Q9fpg1 phaseolus a
   17    71.5    26.8     82  5  Q9U6U0                  Q9u6u0 mytilus gal
   18    71.5    26.8     82  5  Q8MVS6                  Q8mvs6 leiurus qui
   19    71.5    26.8     82  5  Q8MVS8                  Q8mvs8 leiurus qui
   20      71    26.6     81  5  Q9Y0A9                  Q9y0a9 mytilus gal
   21    70.5    26.4     71  5  Q8IRD6                  Q8ird6 drosophila
   22    70.5    26.4     72  5  Q9VZQ4                  Q9vzq4 drosophila
   23    68.5    25.7     85  5  Q9Y0B8                  Q9y0b8 mesobuthus
   24      68    25.5     77  10  Q39403                  Q39403 brassica ca
   25    66.5    24.9   1405  11  Q8VHS2                  Q8vhs2 mus musculu
   26      66    24.7     55  5  Q9TWL1                  Q9twl1 centruroide
   27      66    24.7   3704  5  P91904                  P91904 caenorhabdi
   28    65.5    24.5     64  5  O76963                  O76963 orthochirus
   29    65.5    24.5    428  12  Q99H52                  Q99h52 nectria rad
   30      65    24.3     85  5  Q9BLM4                  Q9blm4 androctonus
c  31      65    27.7    206  12  Q9WHF2                  Q9whf2 agrotis seg
   32    64.5    24.2     85  5  Q26292                  Q26292 leiurus qui
   33      64    24.0     63  10  P82759                  P82759 arabidopsis
   34    63.5    23.8    282  11  Q8BMD9                  Q8bmd9 mus musculu
   35    63.5    23.8   2528  13  Q8AXP0                  Q8axp0 cynops pyrr
   36      63    23.6     72  5  Q9BLM0                  Q9blm0 androctonus
   37      63    23.6     77  10  P82762                  P82762 arabidopsis
   38    62.5    23.4     71  5  Q8IRD7                  Q8ird7 drosophila
   39      62    23.2     72  5  Q9BLM2                  Q9blm2 androctonus
   40      62    23.2     78  10  Q945D8                  Q945d8 castanea sa
   41      62    23.2    128  2  Q8RQM0                  Q8rqm0 streptococc
   42      62    23.2   3712  5  Q9VRW0                  Q9vrw0 drosophila
c  43    61.5    26.2    301  10  O49167                  O49167 zea mays (m
   44    61.5    23.0    569  4  Q8NHD4                  Q8nhd4 homo sapien
   45    61.5    23.0   1406  4  Q8WWY0                  Q8wwy0 homo sapien


                                  ALIGNMENTS

RESULT 1
Q17231
ID   Q17231      PRELIMINARY;      PRT;    65 AA.
AC   Q17231;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Insect toxin 2.
GN   IT2.
OS   Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Mesobuthus.
OX   NCBI_TaxID=34649;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Telson;
RA   Zhu X., Zhang T., Zhu Y.;
RT   "Cloning and sequencing of two depressant insect selective neurotoxin
RT   cDNAs from Buthus martensii Karsch.";
RL   Kexue Tongbao 41:1387-1391(1996).
CC   -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
DR   EMBL; X92077; CAA63059.1; -.
DR   HSSP; P01494; 2SN3.
DR   InterPro; IPR003614; Knot1.
```

```
DR   InterPro; IPR001219; Neurotoxin.
DR   InterPro; IPR002061; Scorpion_toxinL.
DR   Pfam; PF00537; toxin_3; 1.
DR   PRINTS; PR00284; TOXIN.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
SQ   SEQUENCE   65 AA;  7285 MW;  F13FCD7153855626 CRC64;

Alignment Scores:
Pred. No.:                0.0104         Length:        65
Score:                    84.50          Matches:       15
Percent Similarity:       55.26%         Conservative:  6
Best Local Similarity:    39.47%         Mismatches:    12
Query Match:              31.65%         Indels:        5
DB:                       5              Gaps:          2

US-09-673 274B-2_COPY_1_132 (1-132) x Q17231 (1-65)

Qy         16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
              ||:||:::||||||                     ||||||  |||||||:::
Db         14 SerCysLeuTrpGly------------AsnGluGlyCysAsnLysGluCysArgAlaTyr 29

Qy         76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
              |||         |||:::|||   ::::::      :::   |||||||:|||||
Db         30 GlyAlaSerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 46

RESULT 2
Q17230
ID   Q17230       PRELIMINARY;      PRT;    65 AA.
AC   Q17230;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Insect-toxin 3.
GN   IT3.
OS   Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Mesobuthus.
OX   NCBI_TaxID=34649;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Telson;
RA   Zhu X., Zhang T., Zhu Y.;
RT   "Cloning and sequencing of two depressant insect selective neurotoxin
RT   cDNAs from Buthus martensii Karsch.";
RL   Kexue Tongbao 41:1387-1391(1996).
CC   -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
DR   EMBL; X92846; CAA63431.1; -.
DR   HSSP; P01491; 1B3C.
DR   InterPro; IPR003614; Knot1.
DR   InterPro; IPR001219; Neurotoxin.
DR   InterPro; IPR002061; Scorpion_toxinL.
DR   Pfam; PF00537; toxin_3; 1.
DR   PRINTS; PR00284; TOXIN.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
SQ   SEQUENCE   65 AA;  7317 MW;  ECDCC391B733EC0A CRC64;

Alignment Scores:
Pred. No.:                0.0142         Length:        65
Score:                    83.50          Matches:       16
Percent Similarity:       55.26%         Conservative:  5
Best Local Similarity:    42.11%         Mismatches:    12
Query Match:              31.27%         Indels:        5
DB:                       5              Gaps:          2

US-09 673-274B-2_COPY_1_132 (1-132) x Q17230 (1-65)

Qy         16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
              ::||||:::||||||                       ||||   |||||||||
Db         14 SerCysLeuTrpGly------------AsnGluGlyCysAsnLysGluCysLysGlyPhe 29

Qy         76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
              |:|         |||:::|||   ::::::      :::   |||||||||||
Db         30 GlyAlaTyrTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 46

RESULT 3
Q8I0K7
ID   Q8I0K7       PRELIMINARY;      PRT;    85 AA.
AC   Q8I0K7;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Depressant scorpion toxin KIM2 precursor.
GN   KIM2.
OS   Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Mesobuthus.
OX   NCBI_TaxID=34649;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22217688; PubMed=12230558;
RA   Peng F., Zeng X.C., He X.H., Pu J., Li W.X., Zhu Z.H., Liu H.;
RT   "Molecular cloning and functional expression of a gene encoding an
RT   antiarrhythmia peptide derived from the scorpion toxin.";
RL   Eur. J. Biochem. 269:4468-4475(2002).
DR   EMBL; AF459791; AAN59782.1; -.
DR   EMBL; AF459792; AAN59783.1; -.
SQ   SEQUENCE   85 AA;  9425 MW;  5076039B9966AA6F CRC64;

Alignment Scores:
Pred. No.:                0.0144         Length:        85
Score:                    83.50          Matches:       16
Percent Similarity:       55.26%         Conservative:  5
Best Local Similarity:    42.11%         Mismatches:    12
Query Match:              31.27%         Indels:        5
DB:                       5              Gaps:          2

US-09-673-274B-2_COPY_1_132 (1-132) x Q8I0K7 (1-85)

Qy         16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
              |||:||:::|||:||                       ||||||  |||||||||
Db         34 SerCysLeuTrpGly------------AsnGluGlyCysAsnLysGluCysLysGlyPhe 49

Qy         76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
              |||        :||:::|||   ::::::      :::   ||||||||||||
Db         50 GlyAlaTyrTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 66

RESULT 4
Q95WX6
ID   Q95WX6       PRELIMINARY;      PRT;    85 AA.
AC   Q95WX6;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Depressant insect neurotoxin BmK ITb precursor.
OS   Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Mesobuthus.
OX   NCBI_TaxID=34649;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Wang C., Ling M., Wang D., Chi C.;
RT   "The gene for depressant insect toxin BmK ITb from Chinese scorpion,
RT   Buthus martensii Karsch.";
RL   Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
DR   EMBL; AF272777; AAF77063.2; -.
DR   InterPro; IPR003614; Knot1.
DR   InterPro; IPR002061; Scorpion_toxinL.
DR   Pfam; PF00537; toxin_3; 1.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Neurotoxin; Signal.
```

```
FT   SIGNAL        1      21       POTENTIAL.
FT   CHAIN        22      82       DEPRESSANT INSECT NEUROTOXIN BMK ITB.
SQ   SEQUENCE   85 AA;  9330 MW;  80CDD3914956DE8C CRC64;

Alignment Scores:
Pred. No.:               0.0268         Length:        85
Score:                   81.50          Matches:       15
Percent Similarity:      55.26%         Conservative:  6
Best Local Similarity:   39.47%         Mismatches:    12
Query Match:             30.52%         Indels:        5
DB:                      5              Gaps:          2

US-09-673-274B-2_COPY_1_132 (1-132) x Q95WX6 (1-85)

Qy         16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
              ||||||:::||||||                     ||||||   |||||:|||
Db         34 SerCysLeuTrpGly------------AsnGluGlyCysAsnLysGluCysLysAlaPhe 49

Qy         76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
              |||         |||:::|||   ::::::      :::   |||:|||||:::
Db         50 GlyAlaTyrTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGln 66

RESULT 5
Q9XY87
ID   Q9XY87       PRELIMINARY;      PRT;    85 AA.
AC   Q9XY87;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Neurotoxin AEP precursor.
OS   Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Mesobuthus.
OX   NCBI_TaxID=34649;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ling M.H., Wang C.G., Wang D.C., Chi C.W.;
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
DR   EMBL; AF064821; AAD31592.1; -.
DR   HSSP; P01494; 2SN3.
DR   InterPro; IPR003614; Knot1.
DR   InterPro; IPR002061; Scorpion_toxinL.
DR   Pfam; PF00537; toxin_3; 1.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Neurotoxin; Signal.
FT   SIGNAL        1      21       POTENTIAL.
FT   CHAIN        22      85       NEUROTOXIN AEP.
SQ   SEQUENCE   85 AA;  9312 MW;  CCE46711BAF21DAC CRC64;

Alignment Scores:
Pred. No.:               0.0268         Length:        85
Score:                   81.50          Matches:       14
Percent Similarity:      55.26%         Conservative:  7
Best Local Similarity:   36.84%         Mismatches:    12
Query Match:             30.52%         Indels:        5
DB:                      5              Gaps:          2

US-09-673-274B-2_COPY_1_132 (1-132) x Q9XY87 (1-85)

Qy         16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
              ||||||:::||||||                     ||||||   ||||||:::
Db         34 SerCysLeuTrpGly------------AsnGluGlyCysAsnLysGluCysArgAlaTyr 49

Qy         76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
              |||         |||:::|||   ::::::      :::   ||:|||'||:::
Db         50 GlyAlaSerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGln 66

RESULT 6
Q8T3T0
ID   Q8T3T0       PRELIMINARY;      PRT;    85 AA.
AC   Q8T3T0;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Depressant insect toxin BmK ITa1 precursor.
OS   Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Mesobuthus.
OX   NCBI_TaxID=34649;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Liu Z., Chi C., Wu X.;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
DR   EMBL; AY090782; AAM09097.1; -.
DR   InterPro; IPR003614; Knot1.
DR   InterPro; IPR002061; Scorpion_toxinL.
DR   Pfam; PF00537; toxin_3; 1.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Signal.
FT   SIGNAL        1      21       POTENTIAL.
SQ   SEQUENCE   85 AA;  9214 MW;  6CF567CB6032001A CRC64;

Alignment Scores:
Pred. No.:               0.0366         Length:        85
Score:                   80.50          Matches:       15
Percent Similarity:      52.63%         Conservative:  5
Best Local Similarity:   39.47%         Mismatches:    13
Query Match:             30.15%         Indels:        5
DB:                      5              Gaps:          2

US-09-673-274B-2_COPY_1_132 (1-132) x Q8T3T0 (1-85)

Qy         16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
              ||||||:::||||:|                     ||||||   |||||'
Db         34 SerCysLeuTrpGly------------AsnGluGlyCysAsnLysGluCysGlyAlaTyr 49

Qy         76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
              |.|         ||':::||   ::::::      :::   |||||'||||':
Db         50 GlyAlaSerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 66

RESULT 7
P82761
ID   P82761       PRELIMINARY;      PRT;    91 AA.
AC   P82761;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Hypothetical protein LCR46 precursor.
GN   LCR46.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=99397451; PubMed=10470850;
RA   Kaneko T., Katoh T., Sato S., Nakamura Y., Asamizu E., Kotani H.,
RA   Miyajima N., Tabata S.;
RT   "Structural analysis of Arabidopsis thaliana chromosome 5. IX.
RT   Sequence features of the regions of 1,011,550 bp covered by seventeen
RT   P1 and TAC clones.";
RL   DNA Res. 6:183-195(1999).
RN   [2]
RP   IDENTIFICATION.
RX   PubMed=11437247;
RA   Vanoosthuyse V., Miege C., Dumas C., Cock J.M.;
RT   "Two large Arabidopsis thaliana gene families are homologous to the
RT   Brassica gene superfamily that encodes pollen coat proteins and the
```

```
RT   male component of the self-incompatibility response.";
RL   Plant Mol. Biol. 46:17-34(2001).
DR   EMBL; AB017061; -; NOT_ANNOTATED_CDS.
KW   Hypothetical protein; Signal.
FT   SIGNAL         1     27       POTENTIAL.
FT   CHAIN         28     91       HYPOTHETICAL PROTEIN LCR46.
SQ   SEQUENCE   91 AA;  9503 MW;  10BDEAAE80969F55 CRC64;

Alignment Scores:
Pred. No.:                0.0429        Length:         91
Score:                    80.00         Matches:        14
Percent Similarity:       58.06%        Conservative:   4
Best Local Similarity:    45.16%        Mismatches:     13
Query Match:              29.96%        Indels:         0
DB:                       10            Gaps:           0

US-09-673-274B-2_COPY_1_132 (1-132) x P82761 (1-91)

Qy         16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
              :::||.:::   |.|               |||:::||||::      ||: ||
Db         31 ThrCysLeuProGlyGluCysThrAsnProSerGluCysAsnAlaAlaCysLysSerAsn 50

Qy         76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCT 108
              :|||||:|||||:||   :|'   |:'   :::
Db         51 GlyTyrLysGlyGlyAlaCysValSerMetSer 61

RESULT 8
Q9VZQ5
ID   Q9VZQ5       PRELIMINARY;      PRT;    69 AA.
AC   Q9VZQ5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   CG32274 protein.
GN   CG32268 OR CG10815 OR CG11520.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA   Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA   Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA   Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA   Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA   Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA   Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA   Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA   McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA   Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA   Phouanenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA   Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA   Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT   "Sequencing of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA   Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA   Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA   Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA   Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA   Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA   Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT   "Annotation of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RA   FlyBase;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003477; AAF47764.2; -.
DR   HSSP; P41964; 1MYN.
DR   FlyBase; FBgn0052268; CG32268.
DR   InterPro; IPR002118; Gamma-thionin.
DR   ProDom; PD002594; Gamma-thionin; 1.
SQ   SEQUENCE   69 AA;  7720 MW;  87CC05D254A11AC4 CRC64;

Alignment Scores:
Pred. No.:                0.0677        Length:         69
Score:                    78.50         Matches:        16
Percent Similarity:       51.28%        Conservative:   4
Best Local Similarity:    41.03%        Mismatches:     14
Query Match:              29.40%        Indels:         5
DB:                       5             Gaps:           2

US-09-673-274B-2_COPY_1_132 (1-132) x Q9VZQ5 (1-69)

Qy         13 GGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGG 72
              |||||||||         ||||||   :::         |||         |||:::|||
Db         34 GlySerCys---------AlaValTrpHisArgLysLysCysValAspIleCysGlnArg 50

Qy         73 AGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
                 |||         |||||||||              ::::::  ||||||||||||
Db         51 GluGlyArgThrSerGlyHisCysSer------ProSerLeuLysCysTrpCysGlu 67

RESULT 9
Q8MVS7
ID   Q8MVS7       PRELIMINARY;      PRT;    82 AA.
```

```
AC   Q8MVS7;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Depressant insect toxin 2.
OS   Leiurus quinquestriatus quinquestriatus (Egyptian scorpion).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Leiurus.
OX   NCBI_TaxID=6885;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Zaki T.I., Maruniak J.E.;
RT   "Three polymorphic genes encoding LqqIT2 from the Egyptian scorpion
RT   Leiurus quinquestriatus quinquestriatus.";
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
DR   EMBL; AF474984; AAM74028.1; -.
DR   InterPro; IPR003614; Knot1.
DR   InterPro; IPR002061; Scorpion_toxinL.
DR   Pfam; PF00537; toxin_3; 1.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
SQ   SEQUENCE   82 AA;  9099 MW;  F13D53B18CDECBF0 CRC64;

Alignment Scores:
Pred. No.:               0.127          Length:         82
Score:                   76.50          Matches:        15
Percent Similarity:      55.26%         Conservative:   6
Best Local Similarity:   39.47%         Mismatches:     12
Query Match:             28.65%         Indels:         5
DB:                      5              Gaps:           2

US-09-673-274B-2_COPY_1_132 (1-132) x Q8MVS7 (1-82)

Qy         16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
              ||||||::::::|||                     ||||||   ||||:||||
Db         34 SerCysLeuPheGly------------AsnGluGlyCysAsnLysGluCysLysSerTyr 49

Qy         76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
              |||         |||:::|||   ::::::      :::   ||:||||:||||
Db         50 GlyGlySerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 66

RESULT 10
Q9Y1U3
ID   Q9Y1U3      PRELIMINARY;      PRT;    85 AA.
AC   Q9Y1U3;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Depressant insect neurotoxin precursor.
GN   ITC.
OS   Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Mesobuthus.
OX   NCBI_TaxID=34649;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ling M., Wang C., Wang D., Chi C.;
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
DR   EMBL; AF073899; AAD41648.2; -.
DR   HSSP; P01491; 1B3C.
DR   InterPro; IPR003614; Knot1.
DR   InterPro; IPR002061; Scorpion_toxinL.
DR   Pfam; PF00537; toxin_3; 1.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Neurotoxin; Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22     82       DEPRESSANT INSECT NEUROTOXIN.
SQ   SEQUENCE   85 AA;  9480 MW;  18E02CD791F3B5EB CRC64;

Alignment Scores:
Pred. No.:               0.127          Length:         85
Score:                   76.50          Matches:        14
Percent Similarity:      55.26%         Conservative:   7
Best Local Similarity:   36.84%         Mismatches:     12
Query Match:             28.65%         Indels:         5
DB:                      5              Gaps:           2

US-09-673-274B-2_COPY_1_132 (1-132) x Q9Y1U3 (1-85)

Qy         16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
              ||::||::: |||:|                     |||:::      |||||::::
Db         34 SerCysLeuTrpGly------------AsnAspPheCysAspLysValCysLysLysSer 49

Qy         76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
              ||.         |:::::|||   ::::::      :::   |||||:|||||:
Db         50 GlyGlySerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 66

RESULT 11
Q9VZR2
ID   Q9VZR2      PRELIMINARY;      PRT;    69 AA.
AC   Q9VZR2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   CG10812 protein (GH09576p).
GN   CG10812.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
```

```
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA   George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA   Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA   Patel S., Phouanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA   Celniker S.;
RL   Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003477; AAF47757.1; -.
DR   EMBL; AY118762; AAM50622.1; -.
DR   HSSP; P41964; 1MYN.
DR   FlyBase; FBgn0035434; CG10812.
DR   InterPro; IPR002118; Gamma-thionin.
DR   ProDom; PD002594; Gamma-thionin; 1.
SQ   SEQUENCE   69 AA;  7656 MW;  72555FD8B1DCC6F7 CRC64;

Alignment Scores:
Pred. No.:                 0.236          Length:         69
Score:                     74.50          Matches:        15
Percent Similarity:        42.50%         Conservative:   2
Best Local Similarity:     37.50%         Mismatches:     16
Query Match:               27.90%         Indels:         7
DB:                        5              Gaps:           3

US-09-673-274B-2_COPY_1_132 (1-132) x Q9VZR2 (1-69)

Qy          13 GGTTCCTGC---GTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAG 69
                  |||'   |||   |||||                               |||          ||||||
Db          34 GlyProCysAlaValTrpAsp------------AsnGluThrCysArgArgValCysLys 49

Qy          70 AGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
                     |||            |||||||:|           ::::::   ||||||||||||
Db          50 GluGluGlyArgSerSerGlyHisCysSer----- ProSerLeuLysCysTrpCysGlu 67

RESULT 12
Q8WRY4
ID   Q8WRY4      PRELIMINARY;      PRT;    84 AA.
AC   Q8WRY4;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Na+-channel modifying toxin precursor.
GN   CLL9.
OS   Centruroides limpidus limpidus (Mexican scorpion).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Centruroides.
OX   NCBI_TaxID=29941;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Corona M., Coronas F., Merino E., Becerril B., Anturez S.R.,
RA   Garcia D.E., Possani L.D.;
RT   "Cll9, a novel class of peptide found in scorpion venom which
RT   particularly affects Na+-channels of rat peripheral ganglia
RT   preparations.";
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
DR   EMBL; AF387171; AAL57291.1; -.
DR   InterPro; IPR003614; Knot1.
DR   InterPro; IPR002061; Scorpion_toxinL.
DR   Pfam; PF00537; toxin_3; 1.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Signal.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20     82       NA+-CHANNEL MODIFYING TOXIN.
SQ   SEQUENCE   84 AA;  9198 MW;  21018F12E2ECC26E CRC64;

Alignment Scores:
Pred. No.:                 0.238          Length:         84
Score:                     74.50          Matches:        14
Percent Similarity:        48.72%         Conservative:   5
Best Local Similarity:     35.90%         Mismatches:     15
Query Match:               27.90%         Indels:         5
DB:                        5              Gaps:           2

US-09-673-274B-2_COPY_1_132 (1-132) x Q8WRY4 (1-84)

Qy          19 TGCGTGTGGGGTGCTGTGAACTACACTTCCGAT------TGCAACGGTGAGTGCAAGAGG 72
               |||                  ::::::    |||   |||     |||:::      ||||||:::
Db          31 CysThrLeuAlaCysIleAspLysThrGlyAspLysAsnCysAspArgAsnCysLysLys 50

Qy          73 AGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
                     ||.           |||||||||          |||         |||||||||:::
Db          51 GluGlyGlySerPheGlyHisCysSerTyrSerAla---------CysTrpCysLys 66

RESULT 13
Q9GUA7
ID   Q9GUA7      PRELIMINARY;      PRT;    85 AA.
AC   Q9GUA7;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical 9.4 kDa protein.
OS   Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Mesobuthus.
OX   NCBI_TaxID=34649;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Venom gland;
RA   Li W.-X., Zeng X.-C., Zu S.-Y.;
RT   "A new cDNA sequence encoding a toxin from Buthus martensii Karsch.";
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
DR   EMBL; AF150009; AAG09657.1; -.
DR   HSSP; P17728; 1LQI.
DR   InterPro; IPR003614; Knot1.
DR   InterPro; IPR001219; Neurotoxin.
DR   InterPro; IPR002061; Scorpion_toxinL.
DR   Pfam; PF00537; toxin_3; 1.
DR   PRINTS; PR00284; TOXIN.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   85 AA;  9426 MW;  6EFC463E02E95091 CRC64;

Alignment Scores:
Pred. No.:                 0.278          Length:         85
Score:                     74.00          Matches:        12
Percent Similarity:        64.29%         Conservative:   6
Best Local Similarity:     42.86%         Mismatches:     6
Query Match:               27.72%         Indels:         4
DB:                        5              Gaps:           2

US-09-673-274B-2_COPY_1_132 (1-132) x Q9GUA7 (1-85)

Qy          52 TGCAACGGTGAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGC---------GGTTCC 102
               |||:::    |||||||||::::::|||    :::   |||:::|||         |||
Db          41 CysAspAspGluCysLysLysLysGlyAlaGluSerGlyTyrCysGlnTrpAlaGlyVal 60

Qy         103 TTCGCTAACGTGAACTGCTGGTGC 126
               :::   |||      |||||||||
Db          61 TyrGlyAsnAla---CysTrpCys 67

RESULT 14
Q9VZR3
ID   Q9VZR3      PRELIMINARY;      PRT;   106 AA.
AC   Q9VZR3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
```

```
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   CG10813 protein.
GN   CG32279 OR CG10813.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003477; AAF47756.1; -.
DR   HSSP; P41964; 1MYN.
DR   FlyBase; FBgn0052279; CG32279.
DR   InterPro; IPR002118; Gamma-thionin.
DR   InterPro; IPR003614; Knot1.
DR   ProDom; PD002594; Gamma-thionin; 1.
DR   SMART; SM00505; Knot1; 1.
SQ   SEQUENCE   106 AA;  11598 MW;  A76479D6E44BC50C CRC64;

Alignment Scores:
Pred. No.:              0.328          Length:         106
Score:                  73.50          Matches:        16
Percent Similarity:     44.19%         Conservative:   3
Best Local Similarity:  37.21%         Mismatches:     17
Query Match:            27.53%         Indels:         7
DB:                     5              Gaps:           3

US-09-673-274B-2_COPY_1_132 (1-132) x Q9VZR3 (1-106)

Qy          4 AAGCTTATCGGTTCCTGC---GTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
              |||      |||   |||   ||||||                            |||
Db         68 LysTyrLysGlyProCysAlaValTrpAsp------------AsnGluMetCysArgArg 83

Qy         61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
                 ||||::        |||:::        ||||||||||          ::::::   |||
Db         84 IleCysLysGluGluGlyHisIleSerGlyHisCysSer------ProSerLeuLysCys 101

Qy        121 TGGTGCGAG 129
              ||||||||:
Db        102 TrpCysGlu 104

RESULT 15
Q95P69
ID   Q95P69      PRELIMINARY;      PRT;    85 AA.
AC   Q95P69;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Putative sodium channel toxin BmKT.
OS   Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Mesobuthus.
OX   NCBI_TaxID=34649;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=10; TISSUE=Venom gland;
RX   MEDLINE=20191172; PubMed=10728828;
RA   Zeng X.C., Li W.X., Zhu S.Y., Peng F., Jiang D.H., Yang F.H., Wu K.L.;
RT   "Cloning and characterization of the cDNA sequences of two venom
RT   peptides from Chinese scorpion Buthus martensii Karsch (BmK).";
RL   Toxicon 38:893-899(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=10; TISSUE=Venom gland;
RA   Zeng X.C., Li W.X., Zhu S.Y.;
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
DR   EMBL; AF370023; AAK53809.1; -.
DR   InterPro; IPR003614; Knot1.
DR   InterPro; IPR001219; Neurotoxin.
DR   InterPro; IPR002061; Scorpion_toxinL.
DR   Pfam; PF00537; toxin_3; 1.
DR   PRINTS; PR00284; TOXIN.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Ionic channel.
SQ   SEQUENCE   85 AA;  9412 MW;  2FFC463EC2EE57E1 CRC64;

Alignment Scores:
Pred. No.:              0.518          Length:         85
Score:                  72.00          Matches:        12
Percent Similarity:     60.71%         Conservative:   5
Best Local Similarity:  42.86%         Mismatches:     7
Query Match:            26.97%         Indels:         4
DB:                     5              Gaps:           2

US-09-673-274B-2_COPY_1_132 (1-132) x Q95P69 (1-85)

Qy         52 TGCAACGGTGAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGC---------GGTTCC 102
              |||:::   |||||||:::   |||   :::   |||:::|||          |||
Db         41 CysAspAspGluCysLysLysAsnGlyAlaGluSerGlyTyrCysGlnTrpAlaGlyVal 60

Qy        103 TTCGCTAACGTGAACTGCTGGTGC 126
              :::   |||      ||||||||
Db         61 TyrGlyAsnAla---CysTrpCys 67


Search completed: October 16, 2003, 17:16:13
Job time : 40.0299 secs
```

```
                          GenCore version 5.1.6
               Copyright (c) 1993 - 2003  Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on:          October 16, 2003, 17:12:07 ; Search time 27.4086 Seconds
                                          (without alignments)
                                          1578.850 Million cell updates/sec

Title:           US-09-673-274B-2_COPY_1_132
Perfect score:   267
Sequence:        1 gataagcttatcggttcctg..........tgaactgctggtgcgagact 132

Scoring table:   BLOSUM62
                 Xgapop 10.0 , Xgapext  0.5
                 Ygapop 10.0 , Ygapext  0.5
                 Fgapop  6.0 , Fgapext  7.0
                 Delop   6.0 , Delext   7.0

Searched:        609560 seqs, 163917102 residues

Total number of hits satisfying chosen parameters:      1219120

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09673274/runat_16102003_170630_2743/app_query.fasta_1.654
-DB=Published_Applications_AA -QFMT=fastan -SUFFIX=rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
 THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09673274_@CGN_1_1_106_@runat_16102003_170630_2743
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
 FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :       Published_Applications_AA:*
                 1:  /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
                 2:  /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
                 3:  /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
                 4:  /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
                 5:  /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
                 6:  /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
                 7:  /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
                 8:  /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
                 9:  /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
                 10:  /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
                 11:  /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
                 12:  /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
                 13:  /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
                 14:  /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
                 15:  /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
                 16:  /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
                 17:  /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
                 18:  /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                   SUMMARIES

                     %
Result               Query
  No.    Score  Match Length DB  ID                          Description
-------------------------------------------------------------------------------

      1      232    86.9     44  15  US-10-125-258-119          Sequence 119, App
      2      232    86.9     73  15  US-10-125-258-117          Sequence 117, App
      3     79.5    29.8     65  15  US-10-264-480-13           Sequence 13, Appl
      4     79.5    29.8     66  15  US-10-264-480-11           Sequence 11, Appl
      5     76.5    28.7     61  15  US-10-264-480-9            Sequence 9, Appli
      6       75    28.1     61   9  US-09-829-481-2           Sequence 2, Appli
      7     74.5    27.9     44  12  US-10-180-247-4            Sequence 4, Appli
      8     74.5    27.9     69  12  US-10-180-247-6            Sequence 6, Appli
      9     74.5    27.9     70  12  US-10-180-247-2            Sequence 2, Appli
     10     71.5    26.8     61  15  US-10-264-480-6            Sequence 6, Appli
     11     71.5    26.8     61  15  US-10-264-480-7            Sequence 7, Appli
     12     71.5    26.8     61  15  US-10-264-480-8            Sequence 8, Appli
     13       71    26.6     47  15  US-10-178-213-411          Sequence 411, App
     14       71    26.6     77  15  US-10-178-213-410          Sequence 410, App
     15     70.5    26.4     47  15  US-10-178-213-387          Sequence 387, App
     16     70.5    26.4     65  15  US-10-264-480-14           Sequence 14, Appl
     17     70.5    26.4     84  15  US-10-178-213-386          Sequence 386, App
     18     70.5    26.4     87  15  US-10-264-480-10           Sequence 10, Appl
     19     69.5    26.0     84  15  US-10-264-480-18           Sequence 18, Appl
     20       69    25.8     47  15  US-10-178-213-408          Sequence 408, App
     21       69    25.8     47  15  US-10-178-213-444          Sequence 444, App
     22       69    25.8     77  15  US-10-178-213-407          Sequence 407, App
     23       69    25.8     78  15  US-10-178-213-443          Sequence 443, App
c    24     68.5    29.1    173  15  US-10-156-761-7787         Sequence 7787, Ap
     25       68    25.5     47  15  US-10-178-213-396          Sequence 396, App
     26       68    25.5     76  15  US-10-178-213-395          Sequence 395, App
     27     67.5    25.3     87  15  US-10-264-480-12           Sequence 12, Appl
     28       67    25.1     45  15  US-10-178-213-168          Sequence 168, App
     29       67    25.1     72  15  US-10-178-213-167          Sequence 167, App
     30       67    25.1     86  15  US-10-264-480-22           Sequence 22, Appl
     31       67    25.1     86  15  US-10-264-480-23           Sequence 23, Appl
     32     66.5    24.9     47  15  US-10-178-213-309          Sequence 309, App
     33     66.5    24.9     47  15  US-10-178-213-339          Sequence 339, App
     34     66.5    24.9     58  15  US-10-264-480-4            Sequence 4, Appli
     35     66.5    24.9     62  14  US-10-044-359-21           Sequence 21, Appl
     36     66.5    24.9     66  15  US-10-264-480-15           Sequence 15, Appl
     37     66.5    24.9     77  15  US-10-178-213-308          Sequence 308, App
     38     66.5    24.9     84  15  US-10-178-213-338          Sequence 338, App
     39     65.5    24.5     44  15  US-10-178-213-147          Sequence 147, App
     40     65.5    24.5     58  15  US-10-264-480-3            Sequence 3, Appli
     41     65.5    24.5     71  15  US-10-178-213-146          Sequence 146, App
     42       65    24.3     45  15  US-10-178-213-210          Sequence 210, App
     43       65    24.3     47  15  US-10-178-213-3            Sequence 3, Appli
     44       65    24.3     47  15  US-10-178-213-447          Sequence 447, App
     45       65    24.3     47  15  US-10-178-213-450          Sequence 450, App


                                  ALIGNMENTS


RESULT 1
US-10-125-258-119
; Sequence 119, Application US/10125258
; Publication No. US20030028920A1
; GENERAL INFORMATION:
;  APPLICANT: Altier, Daniel J.
;  APPLICANT:  Herrmann, Rafael
;  APPLICANT:  Lu, Albert L.
;  APPLICANT:  McCutchen, Billy F.
;  APPLICANT:  Presnail, James K.
;  APPLICANT:  Weaver, Janine L.
;  APPLICANT:  Wong, James F. H.
;  TITLE OF INVENTION: Antimicrobial Polypeptides and Their
;  TITLE OF INVENTION:  Uses
;  FILE REFERENCE: 35718/246215
;  CURRENT APPLICATION NUMBER: US/10/125,258
;  CURRENT FILING DATE:  2002-04-18
;  PRIOR APPLICATION NUMBER: 60/285,355
;  PRIOR FILING DATE: 2001-04-20
;  NUMBER OF SEQ ID NOS: 127
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 119
;   LENGTH: 44
;   TYPE: PRT
```

```
;   ORGANISM: Agrotis ipsilon
US-10-125-258-119

Alignment Scores:
Pred. No.:              1.97e-21      Length:         44
Score:                  232.00        Matches:        38
Percent Similarity:     93.02%        Conservative:   2
Best Local Similarity:  88.37%        Mismatches:     3
Query Match:            86.89%        Indels:         0
DB:                     15            Gaps:           0

US 09 673 274B 2_COPY_1_132 (1-132) x US-10-125 258-119 (1-44)

Qy          1 GATAAGCTTATCGGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
              :.|      |||||:.:|||||:.:||||::.|||..:::::||.|.|:::.||:||
Db          1 AspValHisIleGlySerCysValTrpGlyAlaValAspTyrThrSerAsnCysAsnAsn 20

Qy         61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
              ::.||||.:.||||..:.||||:|.||||:..|||..|||..|||:|||:||||.:|||:
Db         21 GluCysLysArgArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnValAsnCys 40

Qy        121 TGGTGCGAG 129
              .:|||||.|
Db         41 TrpCysGlu 43

RESULT 2
US-10-125-258-117
; Sequence 117, Application US/10125258
; Publication No. US20030028920A1
; GENERAL INFORMATION:
;  APPLICANT: Altier, Daniel J.
;  APPLICANT:  Herrmann, Rafael
;  APPLICANT:  Lu, Albert L.
;  APPLICANT:  McCutchen, Billy F.
;  APPLICANT:  Presnail, James K.
;  APPLICANT:  Weaver, Janine L.
;  APPLICANT:  Wong, James F. H.
;  TITLE OF INVENTION: Antimicrobial Polypeptides and Their
;  TITLE OF INVENTION:  Uses
;  FILE REFERENCE: 35718/246215
;  CURRENT APPLICATION NUMBER: US/10/125,258
;  CURRENT FILING DATE:  2002-04-18
;  PRIOR APPLICATION NUMBER: 60/285,355
;  PRIOR FILING DATE: 2001-04-20
;  NUMBER OF SEQ ID NOS: 127
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 117
;   LENGTH: 73
;   TYPE: PRT
;   ORGANISM: Agrotis ipsilon
US-10-125-258-117

Alignment Scores:
Pred. No.:              2.07e-21      Length:         73
Score:                  232.00        Matches:        38
Percent Similarity:     93.02%        Conservative:   2
Best Local Similarity:  88.37%        Mismatches:     3
Query Match:            86.89%        Indels:         0
DB:                     15            Gaps:           0

US-09-673 274B 2_COPY_1_132 (1-132) x US-10-125-258-117 (1-73)

Qy          1 GATAAGCTTATCGGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGT 60
              :||      |||:|||||:|:|||||:|||||:.|||:::::.|||||||:::.|||
Db         30 AspValHisIleGlySerCysValTrpGlyAlaValAspTyrThrSerAsnCysAsnAsn 49

Qy         61 GAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGC 120
              ||||.|||||||||||||||:||||:|||||  |||  :|||  :|||.|||||||||||||
Db         50 GluCysLysArgArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnValAsnCys 69

Qy        121 TGGTGCGAG 129
              |||:|.|||
Db         70 TrpCysGlu 72
```

```
RESULT 3
US-10-264-480-13
; Sequence 13, Application US/10264480
; Publication No. US20030113892A1
; GENERAL INFORMATION:
;  APPLICANT: Hammock, Bruce D.
;  APPLICANT:  Inceoglu, Bora
;  TITLE OF INVENTION: ISOLATED POLYPEPTIDES AND COMPOSITIONS
;  TITLE OF INVENTION:  FROM THE VENOM OF P. TRANSVAALICUS AND METHODS OF USE
;  FILE REFERENCE: UCAL256
;  CURRENT APPLICATION NUMBER: US/10/264,480
;  CURRENT FILING DATE:  2002-10-04
;  PRIOR APPLICATION NUMBER: 60/393,070
;  PRIOR FILING DATE: 2002-06-28
;  PRIOR APPLICATION NUMBER: 60/327,602
;  PRIOR FILING DATE: 2001-10-04
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13
;   LENGTH: 65
;   TYPE: PRT
;   ORGANISM: Centruroides sculpturatus
US-10-264-480-13

Alignment Scores:
Pred. No.:              0.0875        Length:         65
Score:                  79.50         Matches:        17
Percent Similarity:     51.28%        Conservative:   3
Best Local Similarity:  43.59%        Mismatches:     12
Query Match:            29.78%        Indels:         7
DB:                     15            Gaps:           3

US-09-673-274B-2_COPY_1_132 (1-132) x US-10-264-480-13 (1-65)

Qy         19 TGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGGGGT 78
              |.:   .||        :||            ||||||   |||||||||  :::
Db         16 CysPheTrpLeuGlyLysAsn------GluHisCysAsnThrGluCysLysAlaLysAsn 33

Qy         79 TACAAGGGT------GGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
                    ||      ||::::|:|   :::||||||        ||||||||:.||
Db         34 GlnGlyGlySerTyrGlyTyrCysTyrAlaPheAla---------CysTrpCysGlu 49

RESULT 4
US-10-264-480-11
; Sequence 11, Application US/10264480
; Publication No. US20030113892A1
; GENERAL INFORMATION:
;  APPLICANT: Hammock, Bruce D.
;  APPLICANT:  Inceoglu, Bora
;  TITLE OF INVENTION: ISOLATED POLYPEPTIDES AND COMPOSITIONS
;  TITLE OF INVENTION:  FROM THE VENOM OF P. TRANSVAALICUS AND METHODS OF USE
;  FILE REFERENCE: UCAL256
;  CURRENT APPLICATION NUMBER: US/10/264,480
;  CURRENT FILING DATE:  2002-10-04
;  PRIOR APPLICATION NUMBER: 60/393,070
;  PRIOR FILING DATE: 2002-06-28
;  PRIOR APPLICATION NUMBER: 60/327,602
;  PRIOR FILING DATE: 2001-10-04
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
;   LENGTH: 66
;   TYPE: PRT
;   ORGANISM: Centruroides limpidus limpidus
US-10-264-480-11

Alignment Scores:
Pred. No.:              0.0876        Length:         66
Score:                  79.50         Matches:        17
Percent Similarity:     53.85%        Conservative:   4
Best Local Similarity:  43.59%        Mismatches:     11
Query Match:            29.78%        Indels:         7
```

```
DB:                     15             Gaps:          3

US-09-673-274B-2_COPY_1_132 (1-132) x US-10-264-480-11 (1-66)

Qy          19 TGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGGGGT 78
               |||   |||         |||         :::|||:::   |||||||||   :::
Db          16 CysPheTrpLeuGlyLysAsn------GluAsnCysAspLysGluCysLysAlaLysAsn 33

Qy          79 TACAAGGGT------GGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
                     |||      |||:::|||   |||||||||         ||||||||||||
Db          34 GlnGlyGlySerTyrGlyTyrCysTyrSerPheAla---------CysTrpCysGlu 49

 ESULT 5
US-10-264-480-9
  Sequence 9, Application US/10264480
  Publication No. US20030113892A1
  GENERAL INFORMATION:
   APPLICANT: Hammock, Bruce D.
   APPLICANT:  Inceoglu, Bora
;  TITLE OF INVENTION: ISOLATED POLYPEPTIDES AND COMPOSITIONS
   TITLE OF INVENTION:  FROM THE VENOM OF P. TRANSVAALICUS AND METHODS OF USE
   FILE REFERENCE: UCAL256
   CURRENT APPLICATION NUMBER: US/10/264,480
   CURRENT FILING DATE:  2002-10-04
   PRIOR APPLICATION NUMBER: 60/393,070
   PRIOR FILING DATE: 2002-06-28
   PRIOR APPLICATION NUMBER: 60/327,602
   PRIOR FILING DATE: 2001-10-04
   NUMBER OF SEQ ID NOS: 30
   SOFTWARE: FastSEQ for Windows Version 4.0
  SEQ ID NO 9
    LENGTH: 61
    TYPE: PRT
    ORGANISM: Leiurus quinquestriatus quinquestriatus
US-10-264-480-9

Alignment Scores:
Pred. No.:              0.212          Length:        61
Score:                  76.50          Matches:       15
Percent Similarity:     55.26%         Conservative:  6
Best Local Similarity:  39.47%         Mismatches:    12
Query Match:            28.65%         Indels:        5
DB:                     15             Gaps:          2

US-09-673-274B-2_COPY_1_132 (1-132) x US-10-264-480-9 (1-61)

Qy          16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
               ||||||::::::|||                     ||||||   |:.|||:||
Db          13 SerCysLeuPheGly------------AsnGluGlyCysAsnLysGluCysLysSerTyr 28

Qy          76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
               |||         |||:::|||   ::::::      :::   :|||:||||||
Db          29 GlyGlySerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 45

RESULT 6
US-09-829-481-2
; Sequence 2, Application US/09829481
; Patent No. US20020069427A1
; GENERAL INFORMATION:
;  APPLICANT: Presnail, James
;  APPLICANT:  Weng, Zude
;  APPLICANT:  Wong, James
;  TITLE OF INVENTION: Arthropod Defensins
;  FILE REFERENCE: BB1441 US NA
;  CURRENT APPLICATION NUMBER: US/09/829,481
;  CURRENT FILING DATE:  2001-04-10
;  PRIOR APPLICATION NUMBER: 60/197279
;  PRIOR FILING DATE: 2000-04-14
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
;   LENGTH: 61
;   TYPE: PRT
```

```
;   ORGANISM: Scolopendra canidens DS
US-09-829-481-2

Alignment Scores:
Pred. No.:              0.33           Length:        61
Score:                  75.00          Matches:       11
Percent Similarity:     53.57%         Conservative:  4
Best Local Similarity:  39.29%         Mismatches:    13
Query Match:            28.09%         Indels:        0
DB:                     9              Gaps:          0

US-09-673-274B-2_COPY_1_132 (1-132) x US-09-829-481-2 (1-61)

Qy          49 GATTGCAACGGTGAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCT 108
               :::.|||||      |:|:::      .||:::   ||||||:::|||         |:|
Db          34 GluCysAsnArgHisCysArgGlyAsnGlyPheThrGlyGlyTyrCysThrGlyPheLeu 53

Qy         109 AACGTGAACTGCTGGTGCGAGACT 132
                        ||    .|     ::
Db          54 LysPheThrCysLysCysTyrThr 61

RESULT 7
US-10-180-247-4
; Sequence 4, Application US/10180247
; Publication No. US20030167519A1
; GENERAL INFORMATION:
;  APPLICANT: DeRose, Richard
;  APPLICANT:  Freysinnet, Georges
;  APPLICANT:  Hoffman, Jules
;  TITLE OF INVENTION: Chimeric Gene Encoding Drosomycin,
;  TITLE OF INVENTION:  Vector Containing It and Production of Disease-Resistant
;  TITLE OF INVENTION:  Transgenic Plants
;  FILE REFERENCE: A32889-PCT-USA-A-A 072667.0182
;  CURRENT APPLICATION NUMBER: US/10/180,247
;  CURRENT FILING DATE:  2002-06-26
;  PRIOR APPLICATION NUMBER: 09/480,251
;  PRIOR FILING DATE: 2000-01-11
;  PRIOR APPLICATION NUMBER: PCT/FR98/01462
;  PRIOR FILING DATE: 1998-07-08
;  PRIOR APPLICATION NUMBER: FR97/09,115
;  PRIOR FILING DATE: 1997-07-11
;  PRIOR APPLICATION NUMBER: FR97/09,663
;  PRIOR FILING DATE: 1997-07-24
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 44
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-10-180-247-4

Alignment Scores:
Pred. No.:              0.371          Length:        44
Score:                  74.50          Matches:       15
Percent Similarity:     42.50%         Conservative:  2
Best Local Similarity:  37.50%         Mismatches:    16
Query Match:            27.90%         Indels:        7
DB:                     12             Gaps:          3

US-09-673-274B-2_COPY_1_132 (1-132) x US-10-180-247-4 (1-44)

Qy          13 GGTTCCTGC---GTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAG 69
               |||   |'|   :|||||                        |||         ||||:|
Db           9 GlyProCysAlaValTrpAsp------------AsnGluThrCysArgArgValCysLys 24

Qy          70 AGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
                     |'|         .||| :|||            ::::::   ||||:|:||||:
Db          25 GluGluGlyArgSerSerGlyHisCysSer------ProSerLeuLysCysTrpCysGlu 42

RESULT 8
US-10-180-247-6
; Sequence 6, Application US/10180247
; Publication No. US20030167519A1
```

```
; GENERAL INFORMATION:
;  APPLICANT: DeRose, Richard
;  APPLICANT:  Freysinnet, Georges
;  APPLICANT:  Hoffman, Jules
;  TITLE OF INVENTION: Chimeric Gene Encoding Drosomycin,
;  TITLE OF INVENTION:  Vector Containing It and Production of Disease-Resistant
;  TITLE OF INVENTION:  Transgenic Plants
;  FILE REFERENCE: A32889-PCT-USA-A-A 072667.0182
;  CURRENT APPLICATION NUMBER: US/10/180,247
;  CURRENT FILING DATE:  2002-06-26
;  PRIOR APPLICATION NUMBER: 09/480,251
;  PRIOR FILING DATE: 2000-01-11
;  PRIOR APPLICATION NUMBER: PCT/FR98/01462
;  PRIOR FILING DATE: 1998-07-08
;  PRIOR APPLICATION NUMBER: FR97/09,115
;  PRIOR FILING DATE: 1997-07-11
;  PRIOR APPLICATION NUMBER: FR97/09,663
;  PRIOR FILING DATE: 1997-07-24
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 69
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Ubiquitin drosomycin fusion peptide
US-10-180-247-6

Alignment Scores:
Pred. No.:                0.387          Length:         69
Score:                    74.50          Matches:        15
Percent Similarity:       42.50%         Conservative:   2
Best Local Similarity:    37.50%         Mismatches:     16
Query Match:              27.90%         Indels:         7
DB:                       12             Gaps:           3

US-09-673-274B-2_COPY_1_132 (1-132) x US-10-180-247-6 (1-69)

Qy          13 GGTTCCTGC---GTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAG 69
               |||   :||   ::|:||                        :||         ||||::
Db          34 GlyProCysAlaValTrpAsp------------AsnGluThrCysArgArgValCysLys 49

Qy          70 AGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
                     |.|         || | |||:            ::::::   ||||||:|||||
Db          50 GluGluGlyArgSerSerGlyHisCysSer------ProSerLeuLysCysTrpCysGlu 67

RESULT 9
US-10-180-247-2
; Sequence 2, Application US/10180247
; Publication No. US20030167519A1
; GENERAL INFORMATION:
;  APPLICANT: DeRose, Richard
;  APPLICANT:  Freysinnet, Georges
;  APPLICANT:  Hoffman, Jules
;  TITLE OF INVENTION: Chimeric Gene Encoding Drosomycin,
;  TITLE OF INVENTION:  Vector Containing It and Production of Disease-Resistant
;  TITLE OF INVENTION:  Transgenic Plants
;  FILE REFERENCE: A32889-PCT-USA-A-A 072667.0182
;  CURRENT APPLICATION NUMBER: US/10/180,247
;  CURRENT FILING DATE:  2002-06-26
;  PRIOR APPLICATION NUMBER: 09/480,251
;  PRIOR FILING DATE: 2000-01-11
;  PRIOR APPLICATION NUMBER: PCT/FR98/01462
;  PRIOR FILING DATE: 1998-07-08
;  PRIOR APPLICATION NUMBER: FR97/09,115
;  PRIOR FILING DATE: 1997-07-11
;  PRIOR APPLICATION NUMBER: FR97/09,663
;  PRIOR FILING DATE: 1997-07-24
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 70
;   TYPE: PRT
```

```
;   ORGANISM: Drosophila melanogaster
US-10-180-247-2

Alignment Scores:
Pred. No.:                0.388          Length:         70
Score:                    74.50          Matches:        15
Percent Similarity:       42.50%         Conservative:   2
Best Local Similarity:    37.50%         Mismatches:     16
Query Match:              27.90%         Indels:         7
DB:                       12             Gaps:           3

US-09-673-274B-2_COPY_1_132 (1-132) x US-10-180-247-2 (1-70)

Qy          13 GGTTCCTGC---GTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAG 69
               .||   |||   ||||||                        |||         |||:||
Db          35 GlyProCysAlaValTrpAsp--- --------AsnGluThrCysArgArgValCysLys 50

Qy          70 AGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
                     |||         ||: |||:|            ::::::   ||||||||:|||
Db          51 GluGluGlyArgSerSerGlyHisCysSer------ProSerLeuLysCysTrpCysGlu 68

RESULT 10
US-10-264-480-6
; Sequence 6, Application US/10264480
; Publication No. US20030113892A1
; GENERAL INFORMATION:
;  APPLICANT: Hammock, Bruce D.
;  APPLICANT:  Inceoglu, Bora
;  TITLE OF INVENTION: ISOLATED POLYPEPTIDES AND COMPOSITIONS
;  TITLE OF INVENTION:  FROM THE VENOM OF P. TRANSVAALICUS AND METHODS OF USE
;  FILE REFERENCE: UCAL256
;  CURRENT APPLICATION NUMBER: US/10/264,480
;  CURRENT FILING DATE:  2002-10-04
;  PRIOR APPLICATION NUMBER: 60/393,070
;  PRIOR FILING DATE: 2002-06-28
;  PRIOR APPLICATION NUMBER: 60/327,602
;  PRIOR FILING DATE: 2001-10-04
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 61
;   TYPE: PRT
;   ORGANISM: Buthacus arenicola
US-10-264-480-6

Alignment Scores:
Pred. No.:                0.932          Length:         61
Score:                    71.50          Matches:        14
Percent Similarity:       55.26%         Conservative:   7
Best Local Similarity:    36.84%         Mismatches:     12
Query Match:              26.78%         Indels:         5
DB:                       15             Gaps:           2

US-09-673-274B-2_COPY_1_132 (1-132) x US-10-264-480-6 (1-61)

Qy          16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
               |||||::::::||||                     |||:::   |||||||||
Db          13 SerCysLeuPheGly------------AsnGluGlyCysAspLysGluCysLysAlaTyr 28

Qy          76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
               |||         |||:::|||   ::::::      :::   ||||||||||||
Db          29 GlyGlySerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 45

RESULT 11
US-10-264-480-7
; Sequence 7, Application US/10264480
; Publication No. US20030113892A1
; GENERAL INFORMATION:
;  APPLICANT: Hammock, Bruce D.
;  APPLICANT:  Inceoglu, Bora
;  TITLE OF INVENTION: ISOLATED POLYPEPTIDES AND COMPOSITIONS
;  TITLE OF INVENTION:  FROM THE VENOM OF P. TRANSVAALICUS AND METHODS OF USE
;  FILE REFERENCE: UCAL256
```

```
  CURRENT APPLICATION NUMBER: US/10/264,480
  CURRENT FILING DATE:  2002-10-04
  PRIOR APPLICATION NUMBER: 60/393,070
  PRIOR FILING DATE: 2002-06-28
  PRIOR APPLICATION NUMBER: 60/327,602
  PRIOR FILING DATE: 2001-10-04
  NUMBER OF SEQ ID NOS: 30
  SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 7
   LENGTH: 61
   TYPE: PRT
   ORGANISM: Buthus occitanus
S-10-264-480-7

lignment Scores:
red. No.:              0.932          Length:        61
core:                  71.50          Matches:       14
ercent Similarity:     55.26%         Conservative:  7
est Local Similarity:  36.84%         Mismatches:    12
uery Match:            26.78%         Indels:        5
B:                     15             Gaps:          2

S-09-673-274B-2_COPY_1_132 (1-132) x US-10-264-480-7 (1-61)

Qy          16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
               ||||||::::::|||                     |||:::   ||:||||:|
Db          13 SerCysLeuPheGly------------AsnGluGlyCysAspLysGluCysLysAlaTyr 28

Qy          76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
               |||         |||:::|||   ::::::      :::   |||||:||||||
Db          29 GlyGlySerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 45

RESULT 12
S-10-264-480-8
 Sequence 8, Application US/10264480
 Publication No. US20030113892A1
 GENERAL INFORMATION:
  APPLICANT: Hammock, Bruce D.
  APPLICANT:  Inceoglu, Bora
  TITLE OF INVENTION: ISOLATED POLYPEPTIDES AND COMPOSITIONS
  TITLE OF INVENTION:  FROM THE VENOM OF P. TRANSVAALICUS AND METHODS OF USE
  FILE REFERENCE: UCAL256
  CURRENT APPLICATION NUMBER: US/10/264,480
  CURRENT FILING DATE:  2002-10-04
  PRIOR APPLICATION NUMBER: 60/393,070
  PRIOR FILING DATE: 2002-06-28
  PRIOR APPLICATION NUMBER: 60/327,602
  PRIOR FILING DATE: 2001-10-04
  NUMBER OF SEQ ID NOS: 30
  SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 8
   LENGTH: 61
   TYPE: PRT
   ORGANISM: Buthus occitanus
S-10-264-480-8

lignment Scores:
red. No.:              0.932          Length:        61
core:                  71.50          Matches:       14
ercent Similarity:     55.26%         Conservative:  7
est Local Similarity:  36.84%         Mismatches:    12
uery Match:            26.78%         Indels:        5
B:                     15             Gaps:          2

S-09-673-274B-2_COPY_1_132 (1-132) x US-10-264-480-8 (1-61)

Qy          16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
               ||||||::::::|||                     |||:::   ||:||||:|
Db          13 SerCysLeuPheGly------------AsnGluGlyCysAspLysGluCysLysAlaTyr 28

Qy          76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
               |||         |||:::|||   ::::::      :.:   |||||:||||||
Db          29 GlyGlySerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 45
```

```
RESULT 13
US-10-178-213-411
; Sequence 411, Application US/10178213
; Publication No. US20030041348A1
; GENERAL INFORMATION:
;  APPLICANT: Simmons, Carl R.
;  APPLICANT:  Navarro Acevedo, Pedro A.
;  APPLICANT:  Harvell, Leslie
;  APPLICANT:  Cahoon, Rebecca
;  APPLICANT:  McCutchen, Billy Fred
;  APPLICANT:  Lu, Albert
;  APPLICANT:  Herrmann, Rafael
;  APPLICANT:  Wong, James
;  TITLE OF INVENTION: Defensin Polynucleotides and Methods of
;  TITLE OF INVENTION:  Use
;  FILE REFERENCE: 35718/246703
;  CURRENT APPLICATION NUMBER: US/10/178,213
;  CURRENT FILING DATE:  2002-06-21
;  PRIOR APPLICATION NUMBER: 60/300,152
;  PRIOR FILING DATE: 2001-06-22
;  PRIOR APPLICATION NUMBER: 60/300,241
;  PRIOR FILING DATE: 2001-06-22
;  NUMBER OF SEQ ID NOS: 469
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 411
;   LENGTH: 47
;   TYPE: PRT
;   ORGANISM: Tropaeolum majus
US-10-178-213-411

Alignment Scores:
Pred. No.:              1.05           Length:        47
Score:                  71.00          Matches:       14
Percent Similarity:     50.00%         Conservative:  5
Best Local Similarity:  36.84%         Mismatches:    13
Query Match:            26.59%         Indels:        6
DB:                     15             Gaps:          2

US-09-673-274B-2_COPY_1_132 (1-132) x US-10-178-213-411 (1-47)

Qy          13 GGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGG 72
               |:|   |:||||                     ::::::|||          ''||||
Db          12 GlyProCysVal---------------SerLysThrAsnCysAlaSerValCysLysThr 26

Qy          73 AGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGC 126
                   '||:::  |: |||:|:|||      |''   :::       |||::::|
Db          27 GluGlyPheTyrGlyGlyHisCysArgGlyPheArgHis-- ArgCysPheCys 43

RESULT 14
US-10-178-213-410
; Sequence 410, Application US/10178213
; Publication No. US20030041348A1
; GENERAL INFORMATION:
;  APPLICANT: Simmons, Carl R.
;  APPLICANT:  Navarro Acevedo, Pedro A.
;  APPLICANT:  Harvell, Leslie
;  APPLICANT:  Cahoon, Rebecca
;  APPLICANT:  McCutchen, Billy Fred
;  APPLICANT:  Lu, Albert
;  APPLICANT:  Herrmann, Rafael
;  APPLICANT:  Wong, James
;  TITLE OF INVENTION: Defensin Polynucleotides and Methods of
;  TITLE OF INVENTION:  Use
;  FILE REFERENCE: 35718/246703
;  CURRENT APPLICATION NUMBER: US/10/178,213
;  CURRENT FILING DATE:  2002-06-21
;  PRIOR APPLICATION NUMBER: 60/300,152
;  PRIOR FILING DATE: 2001-06-22
;  PRIOR APPLICATION NUMBER: 60/300,241
;  PRIOR FILING DATE: 2001-06-22
;  NUMBER OF SEQ ID NOS: 469
;  SOFTWARE: FastSEQ for Windows Version 4.0
```

```
; SEQ ID NO 410
;   LENGTH: 77
;   TYPE: PRT
;   ORGANISM: Tropaeolum majus
US-10-178-213 410

Alignment Scores:
Pred. No.:                1.1          Length:         77
Score:                    71.00        Matches:        14
Percent Similarity:       50.00%       Conservative:   5
Best Local Similarity:    36.84%       Mismatches:     13
Query Match:              26.59%       Indels:         6
DB:                       15           Gaps:           2

US-09-673-274B-2_COPY_1_132 (1-132) x US-10-178-213-410 (1-77)

Qy         13 GGTTCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGG 72
               ||   ! ||||                      ::::::! |         ||| .|
Db         42 GlyProCysVal---------------SerLysThrAsnCysAlaSerValCysLysThr 56

Qy         73 AGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGC 126
                 |||:::   ||||||||||||      '.|   :::      ||':::!'|
Db         57 GluGlyPheTyrGlyGlyHisCysArgGlyPheArgHis---ArgCysPheCys 73

RESULT 15
US-10-178-213 387
; Sequence 387, Application US/10178213
; Publication No. US20030041348A1
; GENERAL INFORMATION:
;  APPLICANT: Simmons, Carl R.
;  APPLICANT:  Navarro Acevedo, Pedro A.
;  APPLICANT:  Harvell, Leslie
;  APPLICANT:  Cahoon, Rebecca
;  APPLICANT:  McCutchen, Billy Fred
;  APPLICANT:  Lu, Albert
;  APPLICANT:  Herrmann, Rafael
;  APPLICANT:  Wong, James
;  TITLE OF INVENTION: Defensin Polynucleotides and Methods of
;  TITLE OF INVENTION:  Use
;  FILE REFERENCE: 35718/246703
,  CURRENT APPLICATION NUMBER: US/10/178,213
;  CURRENT FILING DATE:  2002-06-21
;  PRIOR APPLICATION NUMBER: 60/300,152
;  PRIOR FILING DATE: 2001-06-22
;  PRIOR APPLICATION NUMBER: 60/300,241
;  PRIOR FILING DATE: 2001-06-22
;  NUMBER OF SEQ ID NOS: 469
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 387
;   LENGTH: 47
;   TYPE: PRT
;   ORGANISM: Glycine max
US-10-178-213 387

Alignment Scores:
Pred. No.:                1.22         Length:         47
Score:                    70.50        Matches:        10
Percent Similarity:       59.26%       Conservative:   6
Best Local Similarity:    37.04%       Mismatches:     10
Query Match:              26.40%       Indels:         1
DB:                       15           Gaps:           1

US-09-673-274B-2_COPY_1_132 (1-132) x US-10-178-213-387 (1-47)

Qy         46 TCCGATTGCAACGGTGAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTC 105
              ::::::|||         |||:::       ||:::   ' |||!:|||||      ||.
Db         18 ThrAsnCysGlySerValCysArgThrGluGlyPheSerGlyGlyHisCysArgGlyPhe 37

Qy        106 GCTAACGTGAACTGCTGGTGC 126
              :::         |||:::|||
Db         38 SerArg---ArgCysPheCys 43
```

```
Search completed: October 16, 2003, 17:34:08
Job time : 28.4086 secs
```

```
                GenCore version 5.1.6
         Copyright (c) 1993 - 2003  Compugen Ltd.

M nucleic - protein search, using frame_plus_n2p model

un on:        October 16, 2003, 17:06:37 ; Search time 13.1561 Seconds
                                             (without alignments)
                                             849.038 Million cell updates/sec

itle:         US-09-673-274B-2_COPY_1_132
erfect score: 267
equence:      1 gataagcttatcggttcctg..........tgaactgctggtgcgagact 132

coring table: BLOSUM62
              Xgapop 10.0 , Xgapext  0.5
              Ygapop 10.0 , Ygapext  0.5
              Fgapop  6.0 , Fgapext  7.0
              Delop   6.0 , Delext   7.0

earched:      328717 seqs, 42310858 residues

otal number of hits satisfying chosen parameters:      657434

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
MODEL=frame+_n2p.model -DEV=xlp
Q=/cgn2_1/USPTO_spool/US09673274/runat_16102003_170628_2580/app_query.fasta_1.654
DB=Issued_Patents_AA -QFMT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0
LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
USER=US09673274_@CGN_1_1_55_@runat_16102003_170628_2580 -NCPU=6 -ICPU=3
NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :    Issued_Patents_AA:*
              1:  /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
              2:  /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
              3:  /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
              4:  /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
              5:  /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
              6:  /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                    SUMMARIES
                  %
Result            Query
  No.    Score    Match Length DB  ID                            Description
-----------------------------------------------------------------------------
     1    80.5    30.1     65  1  US-08-435-040-2               Sequence 2, Appli
     2    80.5    30.1     65  3  US-09-020-216-2               Sequence 2, Appli
     3    76.5    28.7     61  1  US-08-682-485A-25             Sequence 25, Appl
     4    76.5    28.7     61  1  US-08-451-472-4               Sequence 4, Appli
     5    76.5    28.7     61  2  US-08-933-314-25              Sequence 25, Appl
     6    74.5    27.9     44  4  US-09-480-251-4               Sequence 4, Appli
     7    74.5    27.9     69  4  US-09-480-251-6               Sequence 6, Appli
     8    74.5    27.9     70  4  US-09-480-251-2               Sequence 2, Appli
     9      74    27.7     64  4  US-09-403-343B-25             Sequence 25, Appl
c   10      68    28.9    147  4  US-09-252-991A-16810          Sequence 16810, A
c   11      68    28.9    149  4  US-09-252-991A-32270          Sequence 32270, A
    12      67    25.1     60  1  US-08-451-472-6               Sequence 6, Appli
    13      67    25.1     67  1  US 08-435 040-3               Sequence 3, Appli
    14      67    25.1     67  3  US-09-020-216-3               Sequence 3, Appli
    15    64.5    24.2     80  3  US-08-952-383A-14             Sequence 14, Appl
    16    64.5    24.2     80  4  US-08-970-264A-27             Sequence 27, Appl
    17    63.5    23.8     61  1  US-08-682-485A-26             Sequence 26, Appl
    18    63.5    23.8     61  2  US-08-933-314-26              Sequence 26, Appl
c   19      63    26.8    631  4  US-09-252-991A-20063          Sequence 20063, A
    20      62    23.2    314  2  US-08-460-309-19              Sequence 19, Appl
    21      62    23.2    314  2  US-08-125-077-19              Sequence 19, Appl
    22    61.5    23.0     74  4  US-09-442-631-4               Sequence 4, Appli
    23    61.5    23.0     75  1  US-08-289-458-2               Sequence 2, Appli
    24    61.5    23.0     75  2  US-08-761-549-2               Sequence 2, Appli
    25    61.5    23.0     75  3  US-09-127-646-2               Sequence 2, Appli
c   26    61.5    26.2    288  4  US-09 252-991A-31433          Sequence 31433, A
c   27    61.5    26.2    301  3  US-09-231-227-4               Sequence 4, Appli
c   28    61.5    26.2    301  4  US-09-768-585-4               Sequence 4, Appli
c   29      61    26.0    296  4  US-09-252-991A-32162          Sequence 32162, A
    30    60.5    22.7     68  3  US-09-053-021-1               Sequence 1, Appli
    31    60.5    22.7     81  3  US 09 053-021-4               Sequence 4, Appli
    32    60.5    22.7     81  3  US-09-053-021-9               Sequence 9, Appli
    33      59    22.1    359  4  US 09-699-266A-11             Sequence 11, Appl
    34    58.5    21.9     67  3  US-09-053-021-2               Sequence 2, Appli
    35      58    21.7     54  1  US-08-656-318A-1              Sequence 1, Appli
    36      58    21.7     54  2  US-08 956-459-1               Sequence 1, Appli
    37      58    21.7     62  1  US-08 451-472-5               Sequence 5, Appli
c   38    57.5    24.5    109  4  US-09-252-991A-22787          Sequence 22787, A
c   39    57.5    24.5    518  4  US-09-252-991A-25967          Sequence 25967, A
    40    57.5    21.5   2556  1  US-08-185-432-17              Sequence 17, Appl
    41    57.5    21.5   2556  1  US-08-083-590A-20             Sequence 20, Appl
    42    57.5    21.5   2556  3  US-08-532-384-20              Sequence 20, Appl
    43    57.5    21.5   2556  4  US-08-899-232-2               Sequence 2, Appli
c   44      57    24.3    145  4  US-09-252-991A-22334          Sequence 22334, A
c   45      57    24.3    526  4  US-09-252-991A-23462          Sequence 23462, A


                                  ALIGNMENTS


RESULT 1
US-08-435-040-2
; Sequence 2, Application US/08435040
; Patent No. 5756340
;  GENERAL INFORMATION:
;    APPLICANT:  Hammock, Bruce D.
;    APPLICANT:  Herrmann, Rafael
;    APPLICANT:  Moskowitz, Haim
;    TITLE OF INVENTION:  Insect Control With Multiple Toxins
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Majestic, Parsons, Siebert & Hsue
;      STREET:  Four Embarcadero Center, Suite 1450
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  94111-4121
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/435,040
;      FILING DATE:  08-MAY-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Siebert, J. Suzanne
;      REGISTRATION NUMBER:  28,758
;      REFERENCE/DOCKET NUMBER:  2500.078US0
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 362-5556
;      TELEFAX:  (415) 362-5418
;      TELEX:  278638 MGPS
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
```

```
;      LENGTH:  65 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
US-08-435-040-2

Alignment Scores:
Pred. No.:                0.0127        Length:         65
Score:                    80.50         Matches:        17
Percent Similarity:       57.14%        Conservative:   7
Best Local Similarity:    40.48%        Mismatches:     9
Query Match:              30.15%        Indels:         9
DB:                       1             Gaps:           4

US-09-673-274B-2_COPY_1_132 (1-132) x US-08-435-040-2 (1-65)

Qy         16 TCCTGCGTGTGG------GGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAG 69
              :::||||||:::      ||||||   :::|||         ||||||   ||||||
Db         12 AsnCysValTyrThrCysGlyAlaAsnSerTyr---------CysAsnThrGluCysThr 28

Qy         70 AGGAGGGGTTACAAGGGTGGTCACTGC---------GGTTCCTTCGCTAACGTGAACTGC 120
              :::   |||   :::   |||:::|||         |||   :::   |||      |||
Db         29 LysAsnGlyAlaGluSerGlyTyrCysGlnTrpPheGlyLysTyrGlyAsnAla---Cys 47

Qy        121 TGGTGC 126
              ||||||
Db         48 TrpCys 49

RESULT 2
US-09-020-216-2
; Sequence 2, Application US/09020216
; Patent No. 6162430
;  GENERAL INFORMATION:
;    APPLICANT:  Hammock, Bruce D.
;    APPLICANT:  Herrmann, Rafael
;    APPLICANT:  Moskowitz, Haim
;    TITLE OF INVENTION:  Insect Control With Multiple Toxins
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Majestic, Parsons, Siebert & Hsue
;      STREET:  Four Embarcadero Center, Suite 1450
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  94111-4121
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/020,216
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/435,040
;      FILING DATE:  08-MAY-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Siebert, J. Suzanne
;      REGISTRATION NUMBER:  28,758
;      REFERENCE/DOCKET NUMBER:  2500.078US0
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 362-5556
;      TELEFAX:  (415) 362-5418
;      TELEX:  278638 MGPS
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  65 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
```

```
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
US-09-020-216-2

Alignment Scores:
Pred. No.:                0.0127        Length:         65
Score:                    80.50         Matches:        17
Percent Similarity:       57.14%        Conservative:   7
Best Local Similarity:    40.48%        Mismatches:     9
Query Match:              30.15%        Indels:         9
DB:                       3             Gaps:           4

US-09-673-274B-2_COPY_1_132 (1-132) x US-09-020-216-2 (1-65)

Qy         16 TCCTGCGTGTGG------GGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAG 69
              :::||||||:::      ||||||   :::|||         ||||||   ||||||
Db         12 AsnCysValTyrThrCysGlyAlaAsnSerTyr---------CysAsnThrGluCysThr 28

Qy         70 AGGAGGGGTTACAAGGGTGGTCACTGC---------GGTTCCTTCGCTAACGTGAACTGC 120
              :::   |||   :::   |||:::|||         |||   :::   |||      |||
Db         29 LysAsnGlyAlaGluSerGlyTyrCysGlnTrpPheGlyLysTyrGlyAsnAla---Cys 47

Qy        121 TGGTGC 126
              ||||||
Db         48 TrpCys 49

RESULT 3
US-08-682-485A-25
; Sequence 25, Application US/08682485A
; Patent No. 5763568
;  GENERAL INFORMATION:
;    APPLICANT:  ATKINSON, RONALD K
;    APPLICANT:  HOWDEN, MERLIN E.H.
;    APPLICANT:  TYLER, MARGARET I
;    APPLICANT:  VONARX, EDWARD J
;    TITLE OF INVENTION:  Insecticidal Toxins Derived From
;    TITLE OF INVENTION:  Funnel Web (Atrax or Hadronyche Spiders)
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Zeneca, Inc.
;      STREET:  1200 South 47th Street
;      CITY:  Richmond
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/682,485A
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/682,485
;      FILING DATE:  17-JULY-1996
;      APPLICATION NUMBER: AUS/08/256,933
;      FILING DATE:  27-JULY-1994
;      APPLICATION NUMBER:  WO 93/15108
;      FILING DATE:  29-JAN-1993
;      APPLICATION NUMBER:  AU PL0722
;      FILING DATE:  31-JAN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Shaw, Melissa A.
;      REGISTRATION NUMBER:  38,301
;      REFERENCE/DOCKET NUMBER:  PPD 5099/D1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  510-231-1542
;      TELEFAX:  510-231-1112
;  INFORMATION FOR SEQ ID NO:  25:
```

```
    SEQUENCE CHARACTERISTICS:
      LENGTH: 61 amino acids
      TYPE:  amino acid
      TOPOLOGY:  linear
    MOLECULE TYPE:  protein
    HYPOTHETICAL:  NO
    ANTI-SENSE:  NO
    ORIGINAL SOURCE:
      ORGANISM:  Leiurus quinquestriatus quinquestriatus
JS-08-682-485A-25

lignment Scores:
red. No.:               0.0418        Length:          61
core:                   76.50         Matches:         15
ercent Similarity:      55.26%        Conservative:    6
est Local Similarity:   39.47%        Mismatches:      12
uery Match:             28.65%        Indels:          5
B:                      1             Gaps:            2

JS-09-673-274B-2_COPY_1_132 (1-132) x US-08-682-485A-25 (1-61)

y          16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
              ||||||::::::|||                     ||||||   |||||||||
b          13 SerCysLeuPheGly------------AsnGluGlyCysAsnLysGluCysLysSerTyr 28

y          76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
              |||         |||:::|||   ::::::      :::   ||||||||||||
b          29 GlyGlySerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 45

ESULT 4
JS-08-451-472-4
  Sequence 4, Application US/08451472
  Patent No. 5770192
   GENERAL INFORMATION:
     APPLICANT:
     TITLE OF INVENTION:  BIOLOGICAL CONTROL AGENTS
,    NUMBER OF SEQUENCES:  73
,    CORRESPONDENCE ADDRESS:
,      ADDRESSEE:  Nixon & Vanderhye PC
,      STREET:  8th Floor, 1100 No. 5770192th Glebe Road
,      CITY:  Arlington
       STATE:  Virginia
,      COUNTRY:  USA
,      ZIP:  22201-4714
,    COMPUTER READABLE FORM:
,      MEDIUM TYPE:  Floppy disk
,      COMPUTER:  IBM PC compatible
,      OPERATING SYSTEM:  PC-DOS/MS-DOS
,      SOFTWARE:  PatentIn Release #1.0, Version #1.25
,    CURRENT APPLICATION DATA:
,      APPLICATION NUMBER:  US/08/451,472
,      FILING DATE:
,      CLASSIFICATION:  435
,    PRIOR APPLICATION DATA:
,      APPLICATION NUMBER:  US 08/117,125
,      FILING DATE:  24-NOV-1993
,    ATTORNEY/AGENT INFORMATION:
,      NAME:  Mary J Wilson
,      REGISTRATION NUMBER:  32,955
;      REFERENCE/DOCKET NUMBER:  117-187
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 816 4000
;      TELEFAX:  (703) 816 4100
;      TELEX:  200797 NIXN UR
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  61 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-451-472-4

Alignment Scores:
```

```
Pred. No.:              0.0418        Length:          61
Score:                  76.50         Matches:         15
Percent Similarity:     55.26%        Conservative:    6
Best Local Similarity:  39.47%        Mismatches:      12
Query Match:            28.65%        Indels:          5
DB:                     1             Gaps:            2

US-09-673-274B-2_COPY_1_132 (1-132) x US-08-451-472-4 (1-61)

Qy          16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
               ||||||::::::|||                     ||||||   |||||||||
Db          13 SerCysLeuPheGly----- --- --AsnGluGlyCysAsnLysGluCysLysSerTyr 28

Qy          76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
               |||         |||:::|||   ::::::      :::   ||||||||||||
Db          29 GlyGlySerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 45

RESULT 5
US-08-933-314-25
; Sequence 25, Application US/08933314
; Patent No. 5959182
;  GENERAL INFORMATION:
;    APPLICANT:  ATKINSON, RONALD K
;    APPLICANT:  HOWDEN, MERLIN E.H.
;    APPLICANT:  TYLER, MARGARET I
;    APPLICANT:  VONARX, EDWARD J
;    TITLE OF INVENTION:  Insecticidal Toxins Derived From
;    TITLE OF INVENTION:  Funnel Web (Atrax or Hadronyche Spiders)
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Zeneca, Inc.
;      STREET:  1200 South 47th Street
;      CITY:  Richmond
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/933,314
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/682,485
;      FILING DATE:  17-JULY-1996
;      APPLICATION NUMBER:  US/08/256,933
;      FILING DATE:  27-JULY-1994
;      APPLICATION NUMBER:  WO 93/15108
;      FILING DATE:  29-JAN-1993
;      APPLICATION NUMBER:  AU PL0722
;      FILING DATE:  31-JAN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Shaw, Melissa A.
;      REGISTRATION NUMBER:  38,321
;      REFERENCE/DOCKET NUMBER:  PPD 5099/D1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  510-231-1542
;      TELEFAX:  510-231-1112
;  INFORMATION FOR SEQ ID NO:  25:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  61 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Leiurus quinquestriatus quinquestriatus
US-08-933-314-25
```

```
Alignment Scores:
Pred. No.:                0.0418        Length:        61
Score:                    76.50         Matches:       15
Percent Similarity:       55.26%        Conservative:  6
Best Local Similarity:    39.47%        Mismatches:    12
Query Match:              28.65%        Indels:        5
DB:                       2             Gaps:          2

US-09-673-274B-2_COPY_1_132 (1-132) x US-08-933-314-25 (1-61)

Qy          16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
                | |||::::::|||                      |||::    :|||:|||
Db          13 SerCysLeuPheGly------------AsnGluGlyCysAsnLysGluCysLysSerTyr 28

Qy          76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
               |||         .'|:::'|'   ::::::      :::   '|||:|'|||:||
Db          29 GlyGlySerTyrGlyTyrCysTrpThrTrp-  GlyLeuAlaCysTrpCysGlu 45

RESULT 6
US-09-480-251 4
; Sequence 4, Application US/09480251
; Patent No. 6465719
; GENERAL INFORMATION:
;  APPLICANT:  DeRose, Richard
;  APPLICANT:  Freyssinet, Georges
;  APPLICANT:  Hoffman, Jules
;  TITLE OF INVENTION:  Chimeric Gene Encoding Drosomycin,
;  TITLE OF INVENTION:  Vector Containing It And Production Of Disease-Resistant
;  TITLE OF INVENTION:  Transgenic Plants
;  FILE REFERENCE:  A32889-PCT-USA-A
;  CURRENT APPLICATION NUMBER: US/09/480,251
;  CURRENT FILING DATE:  2000-01-11
;  PRIOR APPLICATION NUMBER:   PCT/FR98/01462
;  PRIOR FILING DATE:   1998-07-08
;  PRIOR APPLICATION NUMBER:   FRANCE 97/09,115
;  PRIOR FILING DATE:   1997-07-11
;  PRIOR APPLICATION NUMBER:   FRANCE 9709,663
;  PRIOR FILING DATE:   1997-07-24
;  NUMBER OF SEQ ID NOS:   15
;  SOFTWARE:   FastSEQ for Windows Version 3.0
; SEQ ID NO 4
;   LENGTH: 44
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-09-480-251-4

Alignment Scores:
Pred. No.:                0.073         Length:        44
Score:                    74.50         Matches:       15
Percent Similarity:       42.50%        Conservative:  2
Best Local Similarity:    37.50%        Mismatches:    16
Query Match:              27.90%        Indels:        7
DB:                       4             Gaps:          3

US-09-673-274B-2_COPY_1_132 (1-132) x US-09-480-251-4 (1-44)

Qy          13 GGTTCCTGC---GTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAG 69
               |||   |||   |||||:                        |:|         ||||:|
Db           9 GlyProCysAlaValTrpAsp------ - -  AsnGluThrCysArgArgValCysLys 24

Qy          70 AGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
                     :||         |.||||.:            ::::::   ||||:|||.||
Db          25 GluGluGlyArgSerSerGlyHisCysSer------ProSerLeuLysCysTrpCysGlu 42

RESULT 7
US-09-480-251-6
; Sequence 6, Application US/09480251
; Patent No. 6465719
; GENERAL INFORMATION:
;  APPLICANT:  DeRose, Richard
;  APPLICANT:  Freyssinet, Georges
;  APPLICANT:  Hoffman, Jules
;  TITLE OF INVENTION:  Chimeric Gene Encoding Drosomycin,
;  TITLE OF INVENTION:  Vector Containing It And Production Of Disease-Resistant
;  TITLE OF INVENTION:  Transgenic Plants
;  FILE REFERENCE:   A32889-PCT-USA-A
;  CURRENT APPLICATION NUMBER: US/09/480,251
;  CURRENT FILING DATE:  2000-01-11
;  PRIOR APPLICATION NUMBER:   PCT/FR98/01462
;  PRIOR FILING DATE:   1998-07-08
;  PRIOR APPLICATION NUMBER:   FRANCE 97/09,115
;  PRIOR FILING DATE:   1997-07-11
;  PRIOR APPLICATION NUMBER:   FRANCE 9709,663
;  PRIOR FILING DATE:   1997-07-24
;  NUMBER OF SEQ ID NOS:   15
;  SOFTWARE:   FastSEQ for Windows Version 3.0
; SEQ ID NO 6
;   LENGTH: 69
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: The ubiquitin - drosomycin fusion peptide
US-09-480-251-6

Alignment Scores:
Pred. No.:                0.0775        Length:        69
Score:                    74.50         Matches:       15
Percent Similarity:       42.50%        Conservative:  2
Best Local Similarity:    37.50%        Mismatches:    16
Query Match:              27.90%        Indels:        7
DB:                       4             Gaps:          3

US-09-673-274B-2_COPY_1_132 (1-132) x US-09-480-251-6 (1-69)

Qy          13 GGTTCCTGC---GTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAG 69
               ||:   |||   |||||:                        |||         |||||:
Db          34 GlyProCysAlaValTrpAsp------------AsnGluThrCysArgArgValCysLys 49

Qy          70 AGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
                     :|'          |||:.|||            ::::::   |||||||:||||
Db          50 GluGluGlyArgSerSerGlyHisCysSer------ProSerLeuLysCysTrpCysGlu 67

RESULT 8
US-09-480-251-2
; Sequence 2, Application US/09480251
; Patent No. 6465719
; GENERAL INFORMATION:
;  APPLICANT:  DeRose, Richard
;  APPLICANT:  Freyssinet, Georges
;  APPLICANT:  Hoffman, Jules
;  TITLE OF INVENTION:  Chimeric Gene Encoding Drosomycin,
;  TITLE OF INVENTION:  Vector Containing It And Production Of Disease-Resistant
;  TITLE OF INVENTION:  Transgenic Plants
;  FILE REFERENCE:   A32889-PCT-USA-A
;  CURRENT APPLICATION NUMBER: US/09/480,251
;  CURRENT FILING DATE:  2000-01-11
;  PRIOR APPLICATION NUMBER:   PCT/FR98/01462
;  PRIOR FILING DATE:   1998-07-08
;  PRIOR APPLICATION NUMBER:   FRANCE 97/09,115
;  PRIOR FILING DATE:   1997-07-11
;  PRIOR APPLICATION NUMBER:   FRANCE 9709,663
;  PRIOR FILING DATE:   1997-07-24
;  NUMBER OF SEQ ID NOS:   15
;  SOFTWARE:   FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 70
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-09-480-251-2

Alignment Scores:
Pred. No.:                0.0777        Length:        70
Score:                    74.50         Matches:       15
Percent Similarity:       42.50%        Conservative:  2
Best Local Similarity:    37.50%        Mismatches:    16
Query Match:              27.90%        Indels:        7
```

```
B:                        4               Gaps:           3

S-09-673-274B-2_COPY_1_132 (1-132) x US-09-480-251-2 (1-70)

y          13 GGTTCCTGC---GTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAG 69
              |||   |||   ||||||                        |!|         ||||||
b          35 GlyProCysAlaValTrpAsp------------AsnGluThrCysArgArgValCysLys 50

y          70 AGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
                    |||         |||||||||          • ::Y:::   :|||:'|||||:
b          51 GluGluGlyArgSerSerGlyHisCysSer------ProSerLeuLysCysTrpCysGlu 68

ESULT 9
S-09-403-343B-25
 Sequence 25, Application US/09403343B
 Patent No. 6555091
 GENERAL INFORMATION:
  APPLICANT: JOLIVET-REYNAUD, COLETTE
  APPLICANT:  PERRON, HERVE
  APPLICANT:  MANDRAND, BERNARD
  TITLE OF INVENTION: POLYPEPTIDE CAPABLE OF REACTING WITH ANTIBODIES OF
  TITLE OF INVENTION:  PATIENTS SUFFERING FROM MULTIPLE SCLEROSIS AND USES
  FILE REFERENCE: 104574
  CURRENT APPLICATION NUMBER: US/09/403,343B
  CURRENT FILING DATE:  1999-10-18
  PRIOR APPLICATION NUMBER: FR/97/05679
  PRIOR FILING DATE: 1997-04-29
  PRIOR APPLICATION NUMBER: FR/97/16870
  PRIOR FILING DATE: 1997-12-31
  PRIOR APPLICATION NUMBER: PCT/FR98/00870
  PRIOR FILING DATE: 1998-04-29
  NUMBER OF SEQ ID NOS: 34
  SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 25
   LENGTH: 64
   TYPE: PRT
   ORGANISM: MSRV
S-09-403-343B-25

lignment Scores:
red. No.:                0.0892          Length:         64
core:                    74.00           Matches:        12
ercent Similarity:       59.26%          Conservative:   4
est Local Similarity:    44.44%          Mismatches:     9
uery Match:              27.72%          Indels:         2
B:                       4               Gaps:           1

S-09-673-274B-2_COPY_1_132 (1-132) x US-09-403-343B-25 (1-64)

y          52 TGCAACGGTGAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTC------ 105
              |||:::|||   ||||||:::   |||   :||   |||   :|||.:
b          20 CysAspGlyLeuCysLysLysAsnGlyAlaLysSerGlySerCysGlyPheLeuValPro 39

y         106 GCTAACGTGAACTGCTGGTGC 126
              :::   :::   |||||||||
b          40 SerGlyLeuAlaCysTrpCys 46

RESULT 10
US-09-252-991A-16810
, Sequence 16810, Application US/09252991A
, Patent No. 6551795
, GENERAL INFORMATION:
,  APPLICANT:  Marc J. Rubenfield et al.
,  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
,  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
,  FILE REFERENCE:  107196.136
,  CURRENT APPLICATION NUMBER: US/09/252,991A
,  CURRENT FILING DATE:  1999-02-18
,  PRIOR APPLICATION NUMBER:  US 60/074,788
,  PRIOR FILING DATE:  1998-02-18
,  PRIOR APPLICATION NUMBER:  US 60/094,190
,  PRIOR FILING DATE:  1998-07-27
,  NUMBER OF SEQ ID NOS:  33142
```

```
; SEQ ID NO 16810
;   LENGTH: 147
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16810

Alignment Scores:
Pred. No.:                0.605           Length:         147
Score:                    68.00           Matches:        15
Percent Similarity:       59.38%          Conservative:   4
Best Local Similarity:    46.88%          Mismatches:     11
Query Match:              28.94%          Indels:         2
DB:                       4               Gaps:           2

US-09-673-274B-2_COPY_1_132 (1-132) x US-09-252-991A-16810 (1-147)

Qy        106 CGAAGGAACCGCAGTGAC---CACCCTTGTAACCCCTCCTCTTGCACTCACCGTTGCAAT 50
              |:|:|||:.   :::      |||.:|||    '||   :::|||      |:'   ||.
Db         45 ArgArgAsnProAsnProPheHisProCys---ProArgAsnCysAlaGlyArgSerAsn 63

Qy         49 CGGAAGTGTAGTTCACAGCACCCCACACGCAGGAAC 14
              :::   |||         :::|||    ||||::|
Db         64 GlnProCysLeuValProAsnProProArgArgAsn 75

RESULT 11
US-09-252-991A-32270
; Sequence 32270, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 32270
;   LENGTH: 149
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32270

Alignment Scores:
Pred. No.:                0.606           Length:         149
Score:                    68.00           Matches:        17
Percent Similarity:       38.89%          Conservative:   4
Best Local Similarity:    31.48%          Mismatches:     13
Query Match:              28.94%          Indels:         20
DB:                       4               Gaps:           2

US-09-673-274B-2_COPY_1_132 (1-132) x US-09-252-991A-32270 (1-149)

Qy        112 CGTTAGCGAAGGAACCGCAGTGACCACCCT--------------------TGTAACCCC 74
              |..    ||'|||'.|   ':|       ||.                       |||:::|||
Db          7 ArgTrpArgArgAsnSerSerProSerProAlaSerArgArgArgArgArgCysSerPro 26

Qy         73 TCC-------------------------------------TCTTGCACTCACCGTTGC 53
               '|                                              |:|      ||||'.
Db         27 SerArgSerProThrGlyArgTyrThrArgAlaProTrpArgCysCysArgArgArgCys 46

Qy         52 AATCGGAAGTGTAGTTCACAGCACCCCACACGCAGGAACCGA 11
                 |||:::||:       :::   |||||:    :::   |||
Db         47 ArgArgArgCysArgProArgArgProThrSerLysAlaArg 60

RESULT 12
US-08-451-472-6
; Sequence 6, Application US/08451472
; Patent No. 5770192
;  GENERAL INFORMATION:
```

```
;    APPLICANT:
;    TITLE OF INVENTION:  BIOLOGICAL CONTROL AGENTS
;    NUMBER OF SEQUENCES:  73
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Nixon & Vanderhye PC
;      STREET:  8th Floor, 1100 No. 5770192th Glebe Road
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  USA
;      ZIP:  22201-4714
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/451,472
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/117,125
;      FILING DATE:  24-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mary J Wilson
;      REGISTRATION NUMBER:  32,955
;      REFERENCE/DOCKET NUMBER:  117-187
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 816 4000
;      TELEFAX:  (703) 816 4100
;      TELEX:  200797 NIXN UR
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  60 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-451-472-6

Alignment Scores:
Pred. No.:                 0.726          Length:          60
Score:                     67.00          Matches:         13
Percent Similarity:        56.76%         Conservative:    8
Best Local Similarity:     35.14%         Mismatches:      10
Query Match:               25.09%         Indels:          6
DB:                        1              Gaps:            2

US-09-673-274B-2_COPY_1_132 (1-132) x US-08-451-472-6 (1-60)

Qy          19 TGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGGGGT 78
                  |:|:::    :|:                  ::::::|||:::   .|||:|:||      |:|
Db          14 CysLeuIleGly----------------AsnGluCysAspLysGluCysLysAlaTyrGly 28

Qy          79 TACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
                         |||:::|||   ::::::        :::   |||:.|||||||
Db          29 GlySerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 44

RESULT 13
US-08-435-040-3
; Sequence 3, Application US/08435040
; Patent No. 5756340
;  GENERAL INFORMATION:
;    APPLICANT:  Hammock, Bruce D.
;    APPLICANT:  Herrmann, Rafael
;    APPLICANT:  Moskowitz, Haim
;    TITLE OF INVENTION:  Insect Control With Multiple Toxins
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Majestic, Parsons, Siebert & Hsue
;      STREET:  Four Embarcadero Center, Suite 1450
;      CITY:  San Franciso
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  94111-4121
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/435,040
;      FILING DATE:  08-MAY-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Siebert, J. Suzanne
;      REGISTRATION NUMBER:  28,758
;      REFERENCE/DOCKET NUMBER:  2500.078US0
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 362-5556
;      TELEFAX:  (415) 362-5418
;      TELEX:  278638 MGPS
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  67 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
US-08-435-040-3

Alignment Scores:
Pred. No.:                 0.736          Length:          67
Score:                     67.00          Matches:         13
Percent Similarity:        46.15%         Conservative:    5
Best Local Similarity:     33.33%         Mismatches:      19
Query Match:               25.09%         Indels:          2
DB:                        1              Gaps:            1

US-09-673-274B-2_COPY_1 132 (1-132) x US-08-435-040-3 (1-67)

Qy          16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
                  :::|:|||::::                  :::        |||:::        |:||||
Db          12 AsnCysValTyrHisCysPheProGlySerProGlyCysAspThrLeuCysLysGlyAsp 31

Qy          76 GGTTACAAGGGTGGTCACTGCGGT------TCCTTCGCTAACGTGAACTGCTGGTGC 126
                  |||             |||:|||:|:|||                :::   ||||||||:
Db          32 GlyAlaSerSerGlyHisCysGlyPheLysGluGlyHisGlyLeuAlaCysTrpCys 50

RESULT 14
US-09-020-216-3
; Sequence 3, Application US/09020216
; Patent No. 6162430
;  GENERAL INFORMATION:
;    APPLICANT:  Hammock, Bruce D.
;    APPLICANT:  Herrmann, Rafael
;    APPLICANT:  Moskowitz, Haim
;    TITLE OF INVENTION:  Insect Control With Multiple Toxins
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Majestic, Parsons, Siebert & Hsue
;      STREET:  Four Embarcadero Center, Suite 1450
;      CITY:  San Franciso
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  94111-4121
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/020,216
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
```

```
    APPLICATION NUMBER:  US 08/435.040
    FILING DATE:  08-MAY-1995
  ATTORNEY/AGENT INFORMATION:
    NAME:  Siebert, J. Suzanne
    REGISTRATION NUMBER:  28,758
    REFERENCE/DOCKET NUMBER:  2500.078US0
  TELECOMMUNICATION INFORMATION:
    TELEPHONE:  (415) 362-5556
    TELEFAX:  (415) 362-5418
    TELEX:  278638 MGPS
 INFORMATION FOR SEQ ID NO:  3:
  SEQUENCE CHARACTERISTICS:
    LENGTH:  67 amino acids
    TYPE:  amino acid
    TOPOLOGY:  linear
  MOLECULE TYPE:  peptide
  HYPOTHETICAL:  NO
, ANTI-SENSE:  NO
, FRAGMENT TYPE:  N-terminal
US-09-020-216-3

Alignment Scores:
Pred. No.:              0.736          Length:         67
Score:                  67.00          Matches:        13
Percent Similarity:     46.15%         Conservative:   5
Best Local Similarity:  33.33%         Mismatches:     19
Query Match:            25.09%         Indels:         2
DB:                     3              Gaps:           1

US-09-673-274B-2_COPY_1_132 (1-132) x US-09-020-216-3 (1-67)

Qy         16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
              :::||||||:::               :::      |||:::      ||||||
Db         12 AsnCysValTyrHisCysPheProGlySerProGlyCysAspThrLeuCysLysGlyAsp 31

Qy         76 GGTTACAAGGGTGGTCACTGCGGT------TCCTTCGCTAACGTGAACTGCTGGTGC 126
              |||         ||||||||||||                  :::   |||||||||
Db         32 GlyAlaSerSerGlyHisCysGlyPheLysGluGlyHisGlyLeuAlaCysTrpCys 50

RESULT 15
US-08-952-383A-14
  Sequence 14, Application US/08952383A
  Patent No. 6096304
   GENERAL INFORMATION:
     APPLICANT:
     TITLE OF INVENTION:  RECOMBINANT BACULOVIRUS INSECTICIDES
     NUMBER OF SEQUENCES:  14
     COMPUTER READABLE FORM:
       MEDIUM TYPE:  DISKETTE 3.5 INCH
       COMPUTER:  IBM
       OPERATING SYSTEM:  MICROSOFT WINDOWS 3.1
       SOFTWARE:  MICROSOFT WORD FOR WINDOWS 6.0
     CURRENT APPLICATION DATA:
       APPLICATION NUMBER:  US/08/952,383A
       FILING DATE:
,      CLASSIFICATION:  536
,    ATTORNEY/AGENT INFORMATION:
       NAME:  FLOYD, LINDA AXAMETHY
       REGISTRATION NUMBER:  33,692
       REFERENCE/DOCKET NUMBER:  BA-9063-A
,  INFORMATION FOR SEQ ID NO:  14:
,    SEQUENCE CHARACTERISTICS:
,      LENGTH:  80 amino acids
,      TYPE:  amino acid
,      STRANDEDNESS:  single
,      TOPOLOGY:  linear
,    MOLECULE TYPE:  protein
US-08-952-383A-14

Alignment Scores:
Pred. No.:              1.6            Length:         80
Score:                  64.50          Matches:        13
Percent Similarity:     52.63%         Conservative:   7
Best Local Similarity:  34.21%         Mismatches:     13
Query Match:            24.16%         Indels:         5
DB:                     3              Gaps:           2

US-09-673-274B-2_COPY_1_132 (1-132) x US-08-952-383A-14 (1-80)

Qy         16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
              :::;||:::   |||                     |||:::   ||||!·|||
Db         32 AlaCysLeuIleGly---------------AsnGluGlyCysAspLysGluCysLysAlaTyr 47

Qy         76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
              !||         |||:::|'|   ::::::      :::   |.|||||¦'|||
Db         48 GlyGlySerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 64

Search completed: October 16, 2003, 17:13:06
Job time : 14.1561 secs
```

```
                   GenCore version 5.1.6
          Copyright (c) 1993 - 2003  Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on:         October 16, 2003, 17:06:37 ; Search time 20.7741 Seconds
                                             (without alignments)
                                             1564.690 Million cell updates/sec

Title:          US-09-673-274B-2
Perfect score:  324
Sequence:       1 gataagcttatcggttcctg..........gaacgtgtcgacggatccgg 169

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters:      566616

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09673274/runat_16102003_170628_2555/app_query.fasta_1.654
-DB=PIR_76 -QFMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09673274_@CGN_1_1_77_@runat_16102003_170628_2555 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      PIR_76:*
                1:  pir1:*
                2:  pir2:*
                3:  pir3:*
                4:  pir4:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                              SUMMARIES
                    %
Result              Query
  No.     Score     Match Length DB  ID                  Description
------------------------------------------------------------------------------
   1       86       26.5     66  2   A29386              neurotoxin III - s
   2       80.5     24.8     65  2   A59222              neurotoxin X - sco
   3       79.5     24.5     66  2   A55869              crustacean-specifi
   4       79       24.4     85  2   S68906              alpha-toxin Bot XI
   5       78.5     24.2     84  1   NTSR3A              neurotoxin III pre
   6       77.5     23.9     65  1   NTSR4L              neurotoxin IV - Eg
   7       76.5     23.6     61  2   B34123              depressant insect
   8       76       23.5     66  1   NTSR4E              neurotoxin M14 - l
   9       75.5     23.3     65  1   NTSR1C              neurotoxin 1 - bar
  10       75       23.1     64  2   JC1321              neurotoxin IV - Sa
  11       74.5     23.0     70  2   A55824              drosomycin precurs
  12       74       22.8     59  2   C23727              neurotoxin V-5 - b
  13       73       22.5     64  1   NTSR5L              neurotoxin V - Egy
  14       71       21.9   1111  2   T26972              hypothetical prote
  15       70.5     21.8     65  1   NTSR3C              neurotoxin 3 - bar
  16       70.5     21.8     87  2   JN0670              Na+-channel-blocki
  17       69.5     21.5     84  1   NTSR2N              Na+-channel-blocki
  18       68.5     21.1     62  2   D59352              depressant insect-
  19       68.5     21.1     66  1   NTSR9E              neurotoxin M9 - le
  20       68.5     21.1     83  1   NTSRIA              neurotoxin I precu
  21       68       21.0     64  1   NTSR5M              neurotoxin V - sco
  22       68       21.0     77  2   T14395              proteinase inhibit
  23       67.5     20.8     61  2   B59352              depressant insect
  24       67.5     20.8     61  2   C59352              depressant insect-
  25       67.5     20.8     83  1   NTSRI2              neurotoxin I' prec
  26       67.5     20.8     87  2   JN0669              Na+-channel blocki
  27       67       20.7     86  2   JN0671              Na+-channel-blocki
  28       66.5     20.5     65  1   NTSR1B              neurotoxin I - sco
  29       66.5     20.5     66  1   NTSR2C              neurotoxin 2 - bar
c 30       66.5     21.9    873  2   B75514              penicillin-binding
c 31       66.5     21.9   1791  2   T02345              hypothetical prote
  32       66       20.4   3672  2   T23433              hypothetical prote
  33       66       20.4   3704  2   T37316              probable laminin a
  34       65.5     20.2     61  2   A59006              excitatory insect
  35       64.5     19.9     66  2   B23727              neurotoxin V-4 - b
  36       64.5     19.9     85  2   A61616              depressant insect
  37       64.5     19.9    379  2   H96696              protein F1N21.16 [
  38       64       19.8     77  2   S30578              protease inhibitor
  39       64       19.8     78  2   S52634              gamma-thionin - Pe
  40       64       19.8     92  2   S43228              defensin - fruit f
c 41       64       21.1   1017  2   T18488              hypothetical prote
  42       63.5     19.6     85  2   A40472              depressant insect
  43       63       19.4     64  2   A35940              neurotoxin alpha I
  44       63       19.4     85  2   A39306              alpha insect toxin
c 45       63       20.8    965  2   S62935              hypothetical prote

                              ALIGNMENTS

RESULT 1
A29386
neurotoxin III - scorpion (Buthus occitanus)
C;Species: Buthus occitanus mardochei
C;Date: 21-May-1988 #sequence_revision 21-May-1988 #text_change 07-Feb-1997
C;Accession: A29386
R;Vargas, O.; Martin, M.F.; Rochat, H.
Eur. J. Biochem. 162, 589-599, 1987
A;Title: Characterization of six toxins from the venom of the Moroccan scorpion Buthus
A;Reference number: A29386; MUID:87161829; PMID:3104036
A;Accession: A29386
A;Molecule type: protein
A;Residues: 1-66 <VAR>
C;Superfamily: scorpion neurotoxin
C;Keywords: neurotoxin

Alignment Scores:
Pred. No.:              0.0475          Length:         66
Score:                  86.00           Matches:        15
Percent Similarity:     55.00%          Conservative:   7
Best Local Similarity:  37.50%          Mismatches:     16
Query Match:            26.54%          Indels:         2
DB:                     2               Gaps:           1

US-09-673-274B-2 (1-169) x A29386 (1-66)

Qy         16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
                 :::||||||:::                        :::|||   |||:::         |||||    :::
Db         11 AsnCysValTyrHisCysPheProGlySerSerGlyCysAspThrLeuCysLysGluLys 30

Qy         76 GGTTACAAGGGTGGTCACTGCGGTTCCTTC------GCTAACGTGAACTGCTGGTGCGAG 129
                 :||           ||||||||| ||                 :::     |||     |||| |||||:::
Db         31 GlyAlaThrSerGlyHisCysGlyPheLeuProGlySerGlyValAlaCysTrpCysAsp 50

RESULT 2
A59222
neurotoxin X - scorpion (Leiurus quinquestriatus)
C;Species: Leiurus quinquestriatus hebraeus
```

```
C;Date: 31 Dec-2001 #sequence_revision 31-Dec-2001 #text_change 19-Apr-2002
C;Accession: A59222
R;Corzo, G.; Escoubas, P.; Nakajima, T.
submitted to the Protein Sequence Database, April 2000
A;Description: A procedure for the rapid purification of non commercial and new scorpion
A;Reference number: A59222
A;Accession: A59222
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-65 <COR>
C;Superfamily: scorpion neurotoxin
C;Keywords: amidated carboxyl end; neurotoxin; venom
F;13-64,17-37,23 47,27-49/Disulfide bonds: #status predicted
F;65/Modified site: amidated carboxyl end (Arg) #status predicted

Alignment Scores:
Pred. No.:                0.217        Length:        65
Score:                    80.50        Matches:       17
Percent Similarity:       57.14%       Conservative:  7
Best Local Similarity:    40.48%       Mismatches:    9
Query Match:              24.85%       Indels:        9
DB:                       2            Gaps:          4

US-09-673-274B-2 (1-169) x A59222 (1-65)

Qy          16 TCCTGCGTGTGG------GGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAG 69
               :::|::|||:::      |:::||   :::|'.         ! |||.    ||!:||
Db          12 AsnCysValTyrThrCysGlyAlaAsnSerTyr---------CysAsnThrGluCysThr 28

Qy          70 AGGAGGGGTTACAAGGGTGGTCACTGC---------GGTTCCTTCGCTAACGTGAACTGC 120
               :::   .||   :::   |!':::||          |;!    :::    .!      |.
Db          29 LysAsnGlyAlaGluSerGlyTyrCysGlnTrpPheGlyLysTyrGlyAsnAla---Cys 47

Qy         121 TGGTGC 126
               ||!||'
Db          48 TrpCys 49

RESULT 3
A55869
crustacean-specific toxin 1 - scorpion (Centruroides limpidus)
C;Species: Centruroides limpidus limpidus
C;Date: 05 Jan 1996 #sequence_revision 05-Jan-1996 #text_change 17-Mar-1999
C;Accession: A55869
R;Lebreton, F.; Delepierre, M.; Ramirez, A.N.; Balderas, C.; Possani, L.D.
Biochemistry 33, 11135-11149, 1994
A;Title: Primary and NMR three-dimensional structure determination of a novel crustacean
A;Reference number: A55869; MUID:95244424; PMID:7727365
A;Accession: A55869
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-66 <LEB>
C;Superfamily: scorpion neurotoxin
C;Keywords: disulfide bond; toxin
F;12 65,16 41,25-46,29-48/Disulfide bonds: #status experimental

Alignment Scores:
Pred. No.:                0.285        Length:        66
Score:                    79.50        Matches:       17
Percent Similarity:       53.85%       Conservative:  4
Best Local Similarity:    43.59%       Mismatches:    11
Query Match:              24.54%       Indels:        7
DB:                       2            Gaps:          3

US-09-673-274B-2 (1-169) x A55869 (1-66)

Qy          19 TGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGGGGT 78
               |||   :||         ::|         :::|.::::   ||!| ||||   :::
Db          16 CysPheTrpLeuGlyLysAsn------GluAsnCysAspLysGluCysLysAlaLysAsn 33

Qy          79 TACAAGGGT------GGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
                     !||      |||:::|||   ||| '|||.         ;! ||||.||||
Db          34 GlnGlyGlySerTyrGlyTyrCysTyrSerPheAla---- ----CysTrpCysGlu 49

RESULT 4
S68906
alpha-toxin Bot XIV precursor - scorpion (Buthus occitanus)
C;Species: Buthus occitanus tunetanus
C;Date: 23-Jul-1997 #sequence_revision 01-Aug-1997 #text_change 16-Jul-1999
C;Accession: S68906
R;Bouhaouala-Zahar, B.; Ducancel, F.; Zenouaki, I.; ben Khalifa, R.; Borchani, L.; Pel
Eur. J. Biochem. 238, 653-660, 1996
A;Title: A recombinant insect-specific alpha-toxin of Buthus occitanus tunetanus scorp
A;Reference number: S68906; MUID:96300228; PMID:8706664
A;Accession: S68906
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-85 <BOU>
A;Cross-references: GB:X92376; NID:g1041277; PIDN:CAA63120.1; PID:g1041278
C;Superfamily: scorpion neurotoxin
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-85/Product: alpha-toxin #status predicted <MAT>

Alignment Scores:
Pred. No.:                0.318        Length:        85
Score:                    79.00        Matches:       14
Percent Similarity:       50.00%       Conservative:  6
Best Local Similarity:    35.00%       Mismatches:    18
Query Match:              24.38%       Indels:        2
DB:                       2            Gaps:          1

US-09-673-274B-2 (1-169) x S68906 (1-85)

Qy          16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
               :::|||   :::      :::      :::|||   |||:::      ||||||
Db          29 AsnCysAlaTyrHisCysLeuLysIleSerSerGlyCysAspThrLeuCysLysGluAsn 48

Qy          76 GGTTACAAGGGTGGTCACTGCGGT------TCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
               |:|         |'|||: |||.|      |||               |!||||'|!:::
Db          49 GlyAlaThrSerGlyHisCysGlyHisLysSerGlyHisGlySerAlaCysTrpCysLys 68

RESULT 5
NTSR3A
neurotoxin III precursor - Sahara scorpion
N;Alternate names: mammalian neurotoxin II
C;Species: Androctonus australis (Sahara scorpion)
C;Date: 30-Sep-1979 #sequence_revision 04-Oct-1996 #text_change 18-Jun-1999
C;Accession: C34444; A01740
R;Bougis, P.E.; Rochat, H.; Smith, L.A.
J. Biol. Chem. 264, 19259-19265, 1989
A;Title: Precursors of Androctonus australis scorpion neurotoxins. Structures of prec
A;Reference number: A34444; MUID:90037062; PMID:2808423
A;Accession: C34444
A;Molecule type: mRNA
A;Residues: 1-84 <BOU>
A;Cross-references: GB:M27703; GB:J05102; NID:g161142; PIDN:AAA29948.1; PID:g161143
A;Experimental source: clone pcD-634
R;Kopeyan, C.; Martinez, G.; Rochat, H.
Eur. J. Biochem. 94, 609-615, 1979
A;Title: Amino acid sequence of neurotoxin III of the scorpion Androctonus australis
A;Reference number: A01740; MUID:79148414; PMID:428402
A;Accession: A01740
A;Molecule type: protein
A;Residues: 20-26,'N',28-83\<KOP>
C;Superfamily: scorpion neurotoxin
C;Keywords: neurotoxin; venom
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-83/Product: neurotoxin III #status experimental <MAT>
F;31-81,35-53,39-63,43-65/Disulfide bonds: #status predicted

Alignment Scores:
Pred. No.:                0.366        Length:        84
Score:                    78.50        Matches:       15
Percent Similarity:       53.85%       Conservative:  6
Best Local Similarity:    38.46%       Mismatches:    13
Query Match:              24.23%       Indels:        5
DB:                       1            Gaps:          2
```

```
JS-09-673-274B-2 (1-169) x NTSR3A (1-84)

Qy         16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
              :::||||||:::      |||               |||:::|||  |||||||:::
Db         30 AsnCysValTyrHisCysVal---------ProProCysAspGlyLeuCysLysLysAsn 46

Qy         76 GGTTACAAGGGTGGTCACTGCGGTTCCTTC------GCTAACGTGAACTGCTGGTGC 126
              |||   |||   |||   ||||||            :::   :::   |||||||||
Db         47 GlyAlaLysSerGlySerCysGlyPheLeuIleProSerGlyLeuAlaCysTrpCys 65

RESULT 6
NTSR4L
neurotoxin IV - Egyptian scorpion
C;Species: Leiurus quinquestriatus (Egyptian scorpion)
C;Date: 15-Oct-1982 #sequence_revision 27-Nov-1985 #text_change 28-Aug-1998
C;Accession: A01749
R;Kopeyan, C.; Martinez, G.; Rochat, H.
FEBS Lett. 181, 211-217, 1985
A;Title: Primary structure of toxin IV of Leiurus quinquestriatus quinquestriatus.
A;Reference number: A01749
A;Accession: A01749
A;Molecule type: protein
A;Residues: 1-65 <KOP>
A;Note: the source is designated as Leiurus quinquestriatus quinquestriatus
C;Superfamily: scorpion neurotoxin
C;Keywords: blocked carboxyl end; neurotoxin; sodium channel inhibitor; venom
F;13-64,17-37,23-47,27-49/Disulfide bonds: #status predicted
F;65/Modified site: blocked carboxyl end (Arg) (probably amidated) #status experimental

Alignment Scores:
Pred. No.:                0.496         Length:         65
Score:                    77.50         Matches:        16
Percent Similarity:       57.14%        Conservative:   8
Best Local Similarity:    38.10%        Mismatches:     9
Query Match:              23.92%        Indels:         9
DB:                       1             Gaps:           4

US-09-673-274B-2 (1-169) x NTSR4L (1-65)

Qy         16 TCCTGCGTGTGG------GGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAG 69
              :::||||||:::      |||:::   :::|||         ||.|||   ||||||
Db         12 AsnCysValTyrThrCysGlySerAsnSerTyr---------CysAsnThrGluCysThr 28

Qy         70 AGGAGGGGTTACAAGGGTGGTCACTGC---------GGTTCCTTCGCTAACGTGAACTGC 120
              :::   |||   :::   |||:::|||         |||   :::   |||      |||
Db         29 LysAsnGlyAlaGluSerGlyTyrCysGlnTrpLeuGlyLysTyrGlyAsnAla---Cys 47

Qy        121 TGGTGC 126
              ||||||
Db         48 TrpCys 49

RESULT 7
B34123
depressant insect toxin LqqIT2 - scorpion (Leiurus quinquestriatus)
C;Species: Leiurus quinquestriatus quinquestriatus
C;Date: 06-Jul-1990 #sequence_revision 06-Jul-1990 #text_change 02-Sep-2000
C;Accession: B34123; S08268; B38526
R;Zlotkin, E.; Fowler, E.; Eitan, M.; Moyer, M.; Adams, M.E.
Toxicon 28, 170, 1990
A;Title: On the chemistry and action of the depressant insect toxins.
A;Reference number: A34123
A;Accession: B34123
A;Molecule type: protein
A;Residues: 1-61 <ZLO1>
R;Kopeyan, C.; Mansuelle, P.; Sampieri, F.; Brando, T.; Bahraoui, E.M.; Rochat, H.; Gran
FEBS Lett. 261, 423-426, 1990
A;Title: Primary structure of scorpion anti-insect toxins isolated from the venom of Lei
A;Reference number: S08267; MUID:90184494; PMID:2311768
A;Accession: S08268
A;Molecule type: protein
A;Residues: 1-61 <KOP>
R;Zlotkin, E.; Eitan, M.; Bindokas, V.P.; Adams, M.E.; Moyer, M.; Burkhart, W.; Fowler,
```

```
Biochemistry 30, 4814-4821, 1991
A;Title: Functional duality and structural uniqueness of depressant insect-selective ne
A;Reference number: A38526; MUID:91230120; PMID:2029523
A;Accession: B38526
A;Molecule type: protein
A;Residues: 1-61 <ZLO2>
A;Cross-references: CAS:130300-67-3
C;Superfamily: scorpion neurotoxin
C;Keywords: neurotoxin; venom
F;10-60,14-35,21-42,25-44/Disulfide bonds: #status predicted

Alignment Scores:
Pred. No.:                0.658         Length:         61
Score:                    76.50         Matches:        15
Percent Similarity:       55.26%        Conservative:   6
Best Local Similarity:    39.47%        Mismatches:     12
Query Match:              23.61%        Indels:         5
DB:                       2             Gaps:           2

US-09-673-274B-2 (1-169) x B34123 (1-61)

Qy         16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
              ||||| ::::::|||                      .||||:  |||.:|||
Db         13 SerCysLeuPheGly------------AsnGluGlyCysAsnLysGluCysLysSerTyr 28

Qy         76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
              |||         |||:::|||   ::::::      :::   |||||||||||
Db         29 GlyGlySerTyrGlyTyrCysTrpThrTrp   GlyLeuAlaCysTrpCysGlu 45

RESULT 8
NTSR4E
neurotoxin M14 - lesser Asian scorpion
C;Species: Mesobuthus eupeus (lesser Asian scorpion)
C;Date: 30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change 31-Dec-1993
C;Accession: JT0020
R;Volkova, T.M.; Garsia, A.F.; Telezhinskaya, I.N.; Potapenko, N.A.; Grishin, E.V.
Bioorg. Khim. 11, 1445-1456, 1985
A;Title: Study of neurotoxins from the venom of central Asian scorpion Buthus eupeus.
A;Reference number: A94652; MUID:86130759; PMID:4091860
A;Accession: JT0020
A;Molecule type: protein
A;Residues: 1-66 <VOL>
C;Comment: The venom of this scorpion contains three major polypeptide neurotoxins: M9,
C;Superfamily: scorpion neurotoxin
C;Keywords: neurotoxin; venom
F;12-65,16-36,22-46,26-48/Disulfide bonds: #status predicted

Alignment Scores:
Pred. No.:                0.749         Length:         66
Score:                    76.00         Matches:        13
Percent Similarity:       58.62%        Conservative:   4
Best Local Similarity:    44.83%        Mismatches:     8
Query Match:              23.46%        Indels:         4
DB:                       1             Gaps:           2

US-09-673-274B-2 (1-169) x NTSR4E (1-66)

Qy         52 TGCAACGGTGAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGC---------GGTTCC 102
              |||:::   ||||||||::::   |||      |||   :::  ||          .|
Db         22 CysAspSerGluCysLysLysAsnGlyAlaAspGlySerTyrCysGlnTrpLeuGlyArg 41

Qy        103 TTCGCTAACGTGAACTGCTGGTGCGAG 129
              |||   |||      |||||||||:::
Db         42 PheGlyAsnAla---CysTrpCysLys 49

RESULT 9
NTSR1C
neurotoxin 1 - bark scorpion (tentative sequence)
C;Species: Centruroides sculpturatus (bark scorpion)
C;Date: 22-Jun-1981 #sequence_revision 22-Jun-1981 #text_change 31-Mar-2000
C;Accession: A01752
R;Babin, D.R.; Watt, D.D.; Goos, S.M.; Mlejnek, R.V.
Arch. Biochem. Biophys. 164, 694-706, 1974
```

```
A;Title: Amino acid sequences of neurotoxic protein variants from the venom of Centruroi
A;Reference number: A90058; MUID:75163395; PMID:4460885
A;Accession: A01752
A;Molecule type: protein
A;Residues: 1-65 <BAB>
C;Superfamily: scorpion neurotoxin
C;Keywords: neurotoxin; venom
F;12-64,16-41,27-46,29-48/Disulfide bonds: #status predicted

Alignment Scores:
Pred. No.:                0.861        Length:        65
Score:                    75.50        Matches:       14
Percent Similarity:       51.35%       Conservative:  5
Best Local Similarity:    37.84%       Mismatches:    15
Query Match:              23.30%       Indels:        3
DB:                       1            Gaps:          1

US-09-673-274B-2 (1-169) x NTSR1C (1-65)

Qy     19 TGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGGGGT 78
           ||    .|:         |||      :::   |:|      :::   |||   :::|||
Db     16 CysPheTrpLeuGlyLysAsnGluHisAsnThrCysGluCysLysAlaLysAsnGlnGly 35

Qy     79 TACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
                    |||:::'||   :::|||||.         '||'|||||.
Db     36 GlySerTyrGlyTyrCysTyrAlaPheAla---------CysTrpCysGlu 49

RESULT 10
JC1321
neurotoxin IV - Sahara scorpion
C;Species: Androctonus australis (Sahara scorpion)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C;Accession: JC1321
R;Mansuelle, P.; Martin, M.F.; Rochat, H.; Granier, C.
Natural Toxins 1, 61-69, 1992
A;Title: The amino acid sequence of toxin IV from the Androctonus australis scorpion: d
A;Reference number: JC1321; MUID:94221398; PMID:1344902
A;Accession: JC1321
A;Molecule type: protein
A;Residues: 1-64 <MAN>
C;Superfamily: scorpion neurotoxin
C;Keywords: neurotoxin

Alignment Scores:
Pred. No.:                0.99         Length:        64
Score:                    75.00        Matches:       16
Percent Similarity:       51.22%       Conservative:  5
Best Local Similarity:    39.02%       Mismatches:    16
Query Match:              23.15%       Indels:        4
DB:                       2            Gaps:          2

US-09-673-274B-2 (1-169) x JC1321 (1-64)

Qy     52 TGCAACGGTGAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTC------ 105
          .||:::|.|   |.||||:::   .:|   . .     |     |||:
Db     20 CysAspGlyLeuCysLysLysAsnGlyAlaLysSerGlySerCysGlyPheLeuValPro 39

Qy    106 GCTAACGTGAACTGCTGGTGCGAGACTTGAGAGCTCGGCGAGGCGAACGTGTCGACGGAT 165
          :::   :::   |||||||'|         :::|||   ||:               ||:
Db     40 SerGlyLeuAlaCysTrpCysAsn------AspLeuProGluAsnValProIleLysAsp 57

Qy    166 CCG 168
          ||:
Db     58 Pro 58

RESULT 11
A55824
drosomycin precursor - fruit fly (Drosophila melanogaster)
N;Alternate names: cysteine-rich peptide
C;Species: Drosophila melanogaster
C;Date: 23-Mar-1995 #sequence_revision 05-Apr-1995 #text_change 21-Jul-2000
C;Accession: A55824; S38575
R;Fehlbaum, P.; Bulet, P.; Michaut, L.; Lagueux, M.; Broekaert, W.F.; Hetru, C.; Hoffman
```

```
J. Biol. Chem. 269, 33159-33163, 1994
A;Title: Insect immunity. Septic injury of Drosophila induces the synthesis of a pote
A;Reference number: A55824; MUID:95105209; PMID:7806546
A;Accession: A55824
A;Molecule type: mRNA
A;Residues: 1-70 <FEH>
A;Cross-references: EMBL:X75595; NID:g414661; PIDN:CAA53267.1; PID:g414662
A;Note: the mature peptide sequence was confirmed by peptide sequencing
C;Comment: The exact site of signal peptide cleavage (and the requirement, if any, fo
C;Genetics:
A;Gene: FlyBase:Drs
A;Cross-references: FlyBase:FBgn0010381
C;Keywords: disulfide bond; extracellular protein
F;27-70/Product: drosomycin #status experimental <MAT>

Alignment Scores:
Pred. No.:                1.13         Length:        70
Score:                    74.50        Matches:       15
Percent Similarity:       42.50%       Conservative:  2
Best Local Similarity:    37.50%       Mismatches:    16
Query Match:              22.99%       Indels:        7
DB:                       2            Gaps:          3

US-09-673-274B-2 (1-169) x A55824 (1-70)

Qy     13 GGTTCCTGC---GTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAG 69
          ||:   |:    :||||:                        |||         ||:|||
Db     35 GlyProCysAlaValTrpAsp------------AsnGluThrCysArgArgValCysLys 50

Qy     70 AGGAGGGGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
                |.          '|||'|.||            ::::::   ||||||||.|||
Db     51 GluGluGlyArgSerSerGlyHisCysSer------ProSerLeuLysCysTrpCysGlu 68

RESULT 12
C23727
neurotoxin V-5 - bark scorpion
C;Species: Centruroides sculpturatus (bark scorpion)
C;Date: 31-Jan-1992 #sequence_revision 31-Jan-1992 #text_change 07-Feb-1997
C;Accession: C23727
R;David, R.M.; Krishna, N.R.; Watt, D.D.
Toxicon 29, 645-662, 1991
A;Title: Characterization of cationic binding sites of neurotoxins from venom of the
A;Reference number: A23727; MUID:92023284; PMID:1926166
A;Accession: C23727
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-59 <DAV>
C;Superfamily: scorpion neurotoxin
C;Keywords: neurotoxin

Alignment Scores:
Pred. No.:                1.32         Length:        59
Score:                    74.00        Matches:       16
Percent Similarity:       60.53%       Conservative:  7
Best Local Similarity:    42.11%       Mismatches:    7
Query Match:              22.84%       Indels:        8
DB:                       2            Gaps:          3

US-09-673-274B-2 (1-169) x C23727 (1-59)

Qy     16 TCCTGCGTGTGGGGTGCTGTGAACTACACTTCCGATTGCAACGGTGAGTGCAAGAGGAGG 75
          :||||:|||      :||   |||||:         |||:::   :::||||||   :::
Db     14 SerCysVal------AlaAsnAsnTyr---------CysAspAsnGlnCysLysMetLys 28

Qy     76 GGTTACAAGGGTGGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGCGAG 129
                   :'||||||||||   :::         ::::::|||:::||||||
Db     29 LysAlaSerGlyGlyHisCysTyrAla---------MetSerCysTyrCysGlu 43

RESULT 13
NTSR5L
neurotoxin V - Egyptian scorpion
C;Species: Leiurus quinquestriatus (Egyptian scorpion)
C;Date: 30-Apr-1979 #sequence_revision 13-Jul-1981 #text_change 15-Oct-1996
```

```
C;Accession: A01741
R;Kopeyan, C.; Martinez, G.; Rochat, H.
FEBS Lett. 89, 54-58, 1978
A;Title: Amino acid sequence of neurotoxin V from the scorpion Leiurus quinquestriatus q
A;Reference number: A01741; MUID:78191225; PMID:658402
A;Note: L. q. quinquestriatus
A;Accession: A01741
A;Molecule type: protein
A;Residues: 1-64 <KOP>
C;Superfamily: scorpion neurotoxin
C;Keywords: blocked carboxyl end; neurotoxin; venom
F;12-63,16-36,22-46,26-48/Disulfide bonds: #status predicted
F;64/Modified site: blocked carboxyl end (Asn) (probably amidated) #status experimental

Alignment Scores:
Pred. No.:              1.72          Length:         64
Score:                  73.00         Matches:        12
Percent Similarity:     60.71%        Conservative:   5
Best Local Similarity:  42.86%        Mismatches:     7
Query Match:            22.53%        Indels:         4
DB:                     1             Gaps:           2

US-09-673-274B-2 (1-169) x NTSR5L (1-64)

Qy          52 TGCAACGGTGAGTGCAAGAGGAGGGGTTACAAGGGTGGTCACTGC---------GGTTCC 102
               ||||||   |||||||||::::::|||   :::   :||:::|||
Db          22 CysAsnAspGluCysLysLysLysGlyGlyGluSerGlyTyrCysGlnTrpAlaSerPro 41

Qy         103 TTCGCTAACGTGAACTGCTGGTGC 126
               :::   |||      |||||||||
Db          42 TyrGlyAsnAla---CysTrpCys 48

RESULT 14
T26972
hypothetical protein Y47H9C.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000
C;Accession: T26972
R;Harris, B.
submitted to the EMBL Data Library, October 1998
A;Reference number: Z20293
A;Accession: T26972
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1111 <WIL>
A;Cross-references: EMBL:AL032657; PIDN:CAA21739.1; GSPDB:GN00019; CESP:Y47H9C.4
A;Experimental source: clone Y47H9C
C;Genetics:
A;Gene: CESP:Y47H9C.4
A;Map position: 1
A;Introns: 50/2; 84/2; 150/1; 238/3; 342/3; 797/1; 851/1; 947/2; 1017/1; 1083/1
C;Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology

Alignment Scores:
Pred. No.:              2.19          Length:         1111
Score:                  71.00         Matches:        18
Percent Similarity:     38.60%        Conservative:   4
Best Local Similarity:  31.58%        Mismatches:     17
Query Match:            21.91%        Indels:         18
DB:                     2             Gaps:           2

US-09-673-274B-2 (1-169) x T26972 (1-1111)

Qy          49 GATTGC---------------------------AACGGTGAGTGCAAGAGGAGGGGTTAC 81
               ||||||                           :::||||||||||:|      |||:::
Db         293 AspCysLeuAsnAsnGlnAsnCysAspSerSerSerGlyGluCysLysCysIleGlyTrp 312

Qy          82 AAGGGTGGTCACTGC---------------------------GGTTCCTTCGCTAACGTG 114
                  |||   ||||||                           |||
Db         313 ThrGlyLysHisCysAspIleGlyCysSerArgGlyArgPheGlyLeuGlnCysLysGln 332

Qy         115 AACTGCTGGTGCGAGACTTGAGAGCTCGGCGAGGCGAACGTGTCGACGGAT 165
               ||||||   |||         |||      ::::::|||   |||   |||
```

```
Db         333 AsnCysThrCysProGlyLeuGluPheSerAspSerAsnAlaSerCysAsp 349

RESULT 15
NTSR3C
neurotoxin 3 - bark scorpion
C;Species: Centruroides sculpturatus (bark scorpion)
C;Date: 15-Oct-1982 #sequence_revision 15-Oct-1982 #text_change 23-Aug-1996
C;Accession: A90058; A94470; A01754
R;Babin, D.R.; Watt, D.D.; Goos, S.M.; Mlejnek, R.V.
Arch. Biochem. Biophys. 164, 694-706, 1974
A;Title: Amino acid sequences of neurotoxic protein variants from the venom of Centrur
A;Reference number: A90058; MUID:75163395; PMID:4460885
A;Accession: A90058
A;Molecule type: protein
A;Residues: 1-24,'NTC',28-63,'CS' <BAB>
R;Bhown, A.; Mole, J.
unpublished results, cited by Fontecilla-Camps, J.C., et al., Toxicon 20, 1-7, 1982
A;Reference number: A94470
A;Accession: A94470
A;Molecule type: protein
A;Residues: 1-65 <BHO>
R;Fontecilla-Camps, J.C.; Almassy, R.J.; Suddath, F.L.; Bugg, C.E.
Toxicon 20, 1-7, 1982
A;Title: The three-dimensional structure of scorpion neurotoxins.
A;Reference number: A94314; MUID:82200153; PMID:7080025
A;Contents: annotation; X-ray crystallography, 1.8 angstroms; disulfide bonds
A;Note: X-ray crystallographic studies were based on the revised sequence shown
C;Superfamily: scorpion neurotoxin
C;Keywords: neurotoxin; venom
F;12-65,16-41,25-46,29-48/Disulfide bonds: #status experimental

Alignment Scores:
Pred. No.:              3.42          Length:         65
Score:                  70.50         Matches:        15
Percent Similarity:     51.22%        Conservative:   6
Best Local Similarity:  36.59%        Mismatches:     13
Query Match:            21.76%        Indels:         7
DB:                     1             Gaps:           3

US-09-673-274B-2 (1-169) x NTSR3C (1-65)

Qy          19 TGCGTGTGGGGTGCTGTGAACTAC------ACTTCCGATTGCAACGGTGAGTGCAAGAGG 72
               |||   :::|:|   :::                      |||:::    ||||||||::
Db          12 CysLysTyrGlyCysLeuLysLeuGlyGluAsnGluGlyCysAspThrGluCysLysAla 31

Qy          73 AGGGGTTACAAGGGT------GGTCACTGCGGTTCCTTCGCTAACGTGAACTGCTGGTGC 126
               :::         .||       |.|:::||    :::|||||           ||||::|
Db          32 LysAsnGlnGlyGlySerTyrGlyTyrCysTyrAlaPheAla---------CysTrpCys 48

Qy         127 GAG 129
               |:|
Db          49 Glu 49

Search completed: October 16, 2003, 17:11:58
Job time : 23.7741 secs
```